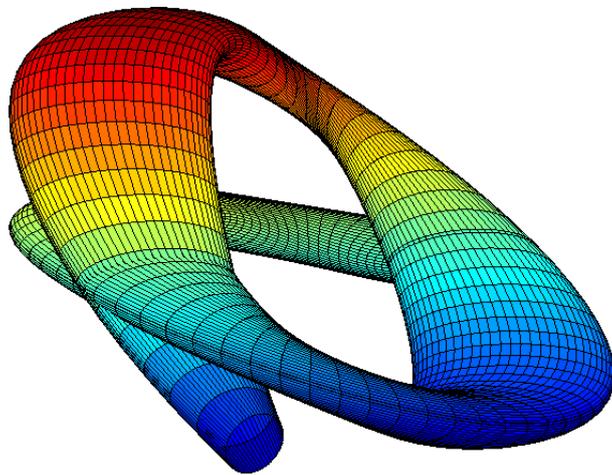


Skriptum zur Vorlesung

# Optimierung



gelesen von

Prof. Dr. Stefan Volkwein

Konstanz, Sommersemester 2015

Martin Gubisch und Konstantin Ott

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>3</b>
<b>2</b>	<b>Optimalitätskriterien</b>	<b>4</b>
2.1	Bedingungen erster und zweiter Ordnung . . . . .	4
2.2	Konvexe Funktionen . . . . .	6
<b>3</b>	<b>Abstiegsverfahren und Schrittweitenstrategien</b>	<b>10</b>
3.1	Abstiegsverfahren . . . . .	11
3.2	Schrittweitenstrategien . . . . .	12
3.2.1	Die Armijo-Regel . . . . .	13
3.2.2	Die WOLFE-POWELL-Regel . . . . .	15
3.2.3	Die strenge Wolfe-Powell-Regel . . . . .	18
3.3	Praktische Aspekte . . . . .	18
3.4	Gradientenverfahren . . . . .	19
<b>4</b>	<b>Verfahren der konjugierten Gradienten</b>	<b>21</b>
4.1	$A$ -konjugierte Richtungen . . . . .	21
4.2	Bestimmung $A$ -konjugierter Richtungen . . . . .	22
<b>5</b>	<b>Newton-Verfahren</b>	<b>24</b>
5.1	Konvergenzraten und das lokale Newton-Verfahren . . . . .	24
5.2	Das lokale Newton-Verfahren in der Optimierung . . . . .	27
5.3	Das inexakte Newton-Verfahren . . . . .	30
5.3.1	Das Newton-CG-Verfahren . . . . .	30
5.4	Das globalisierte Newton-Verfahren . . . . .	31
5.4.1	Abstiegsverfahren . . . . .	31
5.4.2	Die Trust-Region-Methode . . . . .	31
5.5	Quasi-Newton-Verfahren . . . . .	34
	<b>Index</b>	<b>37</b>
	<b>Literaturverzeichnis</b>	<b>38</b>

## 1. Einleitung

Unter einem **endlichdimensionalen Optimierungsproblem** verstehen wir die Aufgabe:

$$\begin{cases} \text{Gegeben sei eine Menge } X \subseteq \mathbb{R}^n \text{ und eine stetige Funktion } f : X \rightarrow \mathbb{R}. \\ \text{Gesucht wird ein } x^* \in X \text{ mit } \forall x \in X : f(x^*) \leq f(x) \end{cases}. \quad (1.1)$$

In Kurznotation:

$$\min f(x) \quad \text{u.d.N.} \quad x \in X \quad \text{bzw.} \quad \min_{x \in X} f(x). \quad (1.2)$$

Ist  $X = \mathbb{R}^n$ , so heißt (1.1) bzw. (1.2) **unrestringiert**, andernfalls **restringiert**.

Im Allgemeinen nennt man  $X$  den **Zulässigkeitsbereich** und  $f$  die **Zielfunktion**.

### Bemerkung 1.1.

Soll  $f$  für  $x \in X$  maximiert werden, so ist dies gleichbedeutend damit, dass  $-f$  u.d.N.  $x \in X$  minimiert wird.  $\diamond$

Die Aufgabenstellung (1.1) erhält ihre Bedeutung dadurch, dass sie ein mathematisches Modell für viele Probleme zum Beispiel aus der Physik, Medizin, Ökonomie und den Ingenieurwissenschaften ist.

Für den Fall, dass  $X \neq \mathbb{R}^n$  ist, lässt sich der Zulässigkeitsbereich sehr häufig in der Form  $X = \Omega_1 \cap \Omega_2 \cap \Omega_3$  schreiben mit

$$\Omega_1 = \{x \in \mathbb{R}^n \mid c_i(x) = 0, i \in I_1\}, \quad \Omega_2 = \{x \in \mathbb{R}^n \mid c_i(x) \leq 0, i \in I_2\}, \quad \Omega_3 = \{x \in \mathbb{R}^n \mid x_i \in \mathbb{Z}, i \in I_3\},$$

gewisse Indexmengen  $I_1, I_2, I_3$  und Abbildungen  $c_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i \in I_1, I_2$ . Die Mengen  $\Omega_1, \Omega_2, \Omega_3$  werden als **Gleichungs-, Ungleichungs-** bzw. **Ganzzahligkeitsrestriktionen** bezeichnet.

Ist  $X$  eine Menge von diskreten Punkten, so spricht man von einem **diskreten** oder **kombinatorischen** Optimierungsproblem, andernfalls von einem **stetigen** Optimierungsproblem.

Ist  $f$  nicht differenzierbar, so nennt man (1.1) ein **nicht-differenzierbares** Optimierungsproblem.

### Definition 1.2.

Sei  $f : X \rightarrow \mathbb{R}$  mit  $X \subseteq \mathbb{R}^n$ . Ein Punkt  $x^* \in X$  heißt

1. **globale Minimalstelle** von  $f$  auf  $X$ , wenn  $f(x^*) \leq f(x)$  für alle  $x \in X$ .  
 $f(x^*)$  heißt dann **globales Minimum**.
2. **strikte globale Minimalstelle** von  $f$  auf  $X$ , wenn  $f(x^*) < f(x)$  für alle  $x \in X \setminus \{x^*\}$ .  
 $f(x^*)$  heißt dann **striktes globales Minimum**.
3. **lokale Minimalstelle** von  $f$  auf  $X$ , wenn es eine Umgebung  $U$  von  $x^*$  gibt, so dass  $f(x^*) \leq f(x)$  für alle  $x \in X \cap U$ .  
 $f(x^*)$  heißt dann **lokales Minimum**.
4. **strikte lokale Minimalstelle** von  $f$  auf  $X$ , wenn es eine Umgebung  $U$  von  $x^*$  gibt mit  $f(x^*) < f(x)$  für alle  $x \in (U \cap X) \setminus \{x^*\}$ .  
 $f(x^*)$  heißt dann **striktes lokales Minimum**.

### Bemerkung 1.3.

Ein Punkt  $x^* \in X$  ist genau dann (**globale, strikte globale, lokale, strikte lokale**) **Maximalstelle** von  $f$  auf  $X$ , wenn  $x^*$  (**globale, strikte globale, lokale, strikte lokale**) Minimalstelle von  $-f$  auf  $X$  ist.  $\diamond$

Die Existenz von (globalen) Lösungen für das Problem (1.2) kann man unter sehr milden Bedingungen zeigen; vergleiche dazu [8, Satz 1.2].

**Satz 1.4.**

Die Funktion  $f : X \rightarrow \mathbb{R}$  sei stetig, und es gebe  $x_0 \in X \subset \mathbb{R}^n$ , so dass die Niveaumenge

$$N_f(x_0) = \{x \in X \mid f(x) \leq f(x_0)\}$$

kompakt ist. Dann besitzt das Problem 1.2 mindestens ein globales Minimum.

**Beweis.**

Es ist klar, dass Punkte, an denen  $f$  sein globales Minimum annimmt, in  $N_f(x_0)$  liegen müssen. Da  $f$  stetig auf der kompakten Menge  $N_f(x_0)$  ist, nimmt  $f$  sein globales Minimum in einem Punkt  $x^* \in N_f(x_0)$  an. Dieser Punkt  $x^*$  ist auch die globale Minimalstelle von  $f$  auf  $X$ .  $\square$

Im Folgenden bezeichne  $\nabla f(x) = \left( \frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right)^T \in \mathbb{R}^n$  den Gradienten von  $f$  in  $x$ .

**Definition 1.5.**

Seien  $X \subseteq \mathbb{R}^n$  eine offene Menge und  $f : X \rightarrow \mathbb{R}$  eine stetig differenzierbare Funktion.

Ein Punkt  $x^* \in X$  heißt **stationärer Punkt** von  $f$ , wenn  $\nabla f(x^*) = 0$  gilt.

**2. Optimalitätskriterien****2.1. Bedingungen erster und zweiter Ordnung**

Wir behandeln unter geeigneten Differenzierbarkeitsannahmen notwendige und hinreichende Bedingungen für lokale Minimalstellen.

**Satz 2.1. (Notwendige Optimalitätsbedingung erster Ordnung)**

Seien  $X \subseteq \mathbb{R}^n$  offen und  $f : X \rightarrow \mathbb{R}$  stetig differenzierbar. Ist  $x^* \in X$  eine lokale Minimalstelle von  $f$  auf  $X$ , so gilt  $\nabla f(x^*) = 0$ , d.h.  $x^*$  ist ein stationärer Punkt.

**Beweis.** (Analysis II)

Sei  $x^* \in X$  eine lokale Minimalstelle von  $f$ , aber  $\nabla f(x^*) \neq 0$ . Dann existiert  $d \in \mathbb{R}^n$  mit  $\nabla f(x^*)^T d < 0$  (z.B.  $d = -\nabla f(x^*)$ ). Da nach Voraussetzung  $f$  stetig differenzierbar ist, existiert die Richtungsableitung  $f'(x^*; d)$  von  $f$  in  $x^*$  in Richtung  $d$ . Es gilt

$$f'(x^*; d) = \lim_{t \searrow 0} \frac{f(x^* + td) - f(x^*)}{t} = \nabla f(x^*)^T d < 0.$$

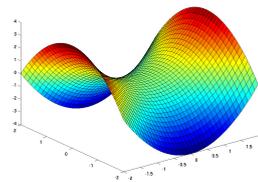
Folglich existiert ein  $t_0 > 0$  mit  $x^* + td \in X$  und  $\frac{1}{t}(f(x^* + td) - f(x^*)) < 0$  für alle  $t \in (0, t_0]$ . Somit ist auch  $f(x^* + td) < f(x^*)$  für alle  $t \in (0, t_0]$ , was einen Widerspruch zur Voraussetzung ergibt.  $\square$

**Bemerkung 2.2.**

Die Bedingungen aus Satz 2.1 sind nicht hinreichend dafür, dass  $x^*$  eine lokale Minimalstelle ist.

Betrachte zum Beispiel die Funktion  $f(x) = x_1^2 - x_2^2$  auf  $\mathbb{R}^2$  und den Punkt  $x^* = (0, 0)$ , dann gilt  $\nabla f(x_1, x_2) = (2x_1, -2x_2)^T$ , also ist  $x^* = (0, 0)$  ein stationärer Punkt von  $f$ , aber für beliebig kleines  $\epsilon > 0$  ist  $f(x_\epsilon) = -\epsilon^2 < 0 = f(x^*)$  mit  $x_\epsilon = (0, \epsilon)$ , d.h. in jeder Umgebung von  $x^*$  existiert ein Punkt mit kleinerem Funktionswert.

Da Satz 2.1 nur Ableitungen bis zur ersten Ordnung verwendet, gibt er eine notwendige Bedingung erster Ordnung an.  $\diamond$



**Fig. 1:** Eine Funktion mit einem **Sattelpunkt**, d.h. einem kritischen Punkt, der kein Optimum ist.

Wir zitieren folgende Abschätzung aus der Spektraltheorie symmetrischer Matrizen:

**Lemma 2.3.**

Sei  $\mathcal{S}_n$  der Vektorraum der symmetrischen  $(n \times n)$ -Matrizen. Für  $A \in \mathcal{S}_n$  sei  $\lambda(A)$  der kleinste Eigenwert von  $A$ . Dann gilt

$$|\lambda(A) - \lambda(B)| \leq \|A - B\|_2 \quad \text{für alle } A, B \in \mathcal{S}_n,$$

wobei  $\|\cdot\|_2$  hier die **Spektralnorm**  $\|A\|_2 = \max\{|\lambda| \mid \lambda \text{ ist Eigenwert von } A\}$  bezeichnet.

Vergleiche dazu das Kapitel zur Störepfindlichkeit bei Eigenwertaufgaben aus [6, Kapitel 11.4].

Mithilfe Lemma 2.3 folgt aus der Stetigkeit der Hessematrix  $\nabla^2 f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right) \in \mathbb{R}^{n \times n}$  von  $f$ : Ist  $\nabla^2 f(x^*)$  positiv definit, dann ist  $\nabla^2 f(x)$  positiv definit in einer Umgebung von  $x^*$ . Eine analoge Folgerung gilt für den Fall, dass  $\nabla^2 f(x^*)$  negativ definit ist.

**Satz 2.4. (Notwendige Optimalitätsbedingung zweiter Ordnung)**

Seien  $X \subseteq \mathbb{R}^n$  offen und  $f : X \rightarrow \mathbb{R}$  zweimal stetig differenzierbar. Ist  $x^* \in X$  eine lokale Minimalstelle von  $f$  auf  $X$ , so ist  $\nabla f(x^*) = 0$  und  $\nabla^2 f(x^*)$  ist positiv semidefinit.

**Beweis.** (Analysis II)

Die Bedingung  $\nabla f(x^*) = 0$  folgt aus Satz 2.1. Sei  $x^*$  eine lokale Minimalstelle von  $f$ , jedoch  $\nabla^2 f(x^*)$  nicht positiv semidefinit. Dann existiert ein  $d \in \mathbb{R}^n$  mit  $d^T \nabla^2 f(x^*) d < 0$ .

Mit Hilfe des Satzes von TAYLOR ergibt sich

$$f(x^* + td) = f(x^*) + \underbrace{t \nabla f(x^*)^T}_{=0} d + \frac{1}{2} t^2 d^T \nabla^2 f(\xi_t) d$$

für kleines  $t > 0$ . Dabei ist  $\xi_t = x^* + \vartheta_t t d$  für  $\vartheta_t \in (0, 1)$ . Aus Lemma 2.3 und der Stetigkeit der zweiten Ableitung von  $f$  folgt die Existenz von  $t_0 > 0$  mit  $d^T \nabla^2 f(\xi_t) d < 0$  für alle  $t \in (0, t_0]$ , wegen  $\nabla f(x^*) = 0$  also  $f(x^* + td) < f(x^*)$  für alle  $t \in (0, t_0]$ , was einen Widerspruch zur Voraussetzung ergibt.  $\square$

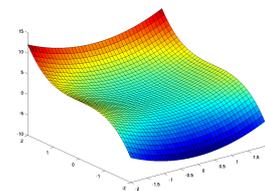
**Bemerkung 2.5.**

Die Bedingungen aus Satz 2.4 sind nicht hinreichend dafür, dass  $x^*$  eine lokale Minimalstelle ist.

Betrachte zum Beispiel die Funktion  $f(x) = x_1^2 + x_2^3$  auf  $\mathbb{R}^2$  und den Punkt  $x^* = (0, 0)^T$ , dann gelten

$$\nabla f(x^*) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \nabla^2 f(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Da Satz 2.4 Ableitungen bis zur zweiten Ordnung verwendet, gibt es eine notwendige Bedingung zweiter Ordnung an.  $\diamond$



**Fig. 2:** Ein Sattelpunkt mit semidefiniter Hessematrix.

Nun kommen wir zu einer hinreichenden Optimalitätsbedingung.

**Satz 2.6. (Hinreichende Optimalitätsbedingung zweiter Ordnung)**

Seien  $X \subseteq \mathbb{R}^n$  offen und  $f : X \rightarrow \mathbb{R}$  zweimal stetig differenzierbar. Gilt  $\nabla f(x^*) = 0$  und ist  $\nabla^2 f(x^*)$  positiv definit, dann ist  $x^*$  eine strikte lokale Minimalstelle von  $f$  auf  $X$ .

**Beweis.** (Analysis II)

Aus der positiven Definitheit der Hessematrix folgt die Existenz eines  $\mu > 0$  mit  $d^T \nabla^2 f(x^*) d \geq \mu d^T d$  für alle  $d \in \mathbb{R}^n$ , wähle etwa  $\mu$  als den kleinsten Eigenwert von  $\nabla^2 f(x^*)$ . Da  $x^*$  ein stationärer Punkt ist, gilt nach dem Satz von TAYLOR für alle  $d \in \mathbb{R}^n$ , die hinreichend nahe bei 0 aber  $d \neq 0$  sind, dass

$$f(x^* + d) = f(x^*) + \underbrace{\nabla f(x^*)^T d}_{=0} + \frac{1}{2} d^T \nabla^2 f(\xi_d) d$$

mit  $\xi_d = x^* + \vartheta_d d$  für  $\vartheta_d \in (0, 1)$  erfüllt ist. Man erhält so mit der CAUCHY-SCHWARZschen Ungleichung

$$\begin{aligned} f(x^* + d) &= f(x^*) + \frac{1}{2} d^T \nabla^2 f(x^*) d + \frac{1}{2} d^T (\nabla^2 f(\xi_d) - \nabla^2 f(x^*)) d \\ &\geq f(x^*) + \frac{1}{2} (\mu - \|\nabla^2 f(\xi_d) - \nabla^2 f(x^*)\|_2) \|d\|_2^2 > f(x^*). \end{aligned}$$

Da  $f$  zweimal stetig differenzierbar ist, gibt es ein  $\delta = \delta(\mu) > 0$ , so dass  $\|\nabla^2 f(\xi_d) - \nabla^2 f(x^*)\|_2 \leq \mu/2$  gilt für alle  $d \in \mathbb{R}^n$  mit  $\|d\|_2 < \delta$ . Dann erhalten wir

$$f(x^* + d) \geq f(x^*) + \frac{\mu}{4} \|d\|_2^2 > f(x^*) \quad \text{für alle } d \neq 0 \text{ mit } \|d\|_2 < \delta,$$

das heißt,  $x^*$  ist eine strikte lokale Minimalstelle von  $f$ . □

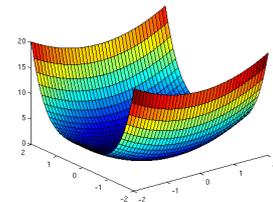
**Bemerkung 2.7.**

Die Bedingungen aus Satz 2.6 sind nicht notwendig dafür, dass  $x^*$  eine lokale Minimalstelle ist.

Betrachte zum Beispiel die Funktion  $f(x) = x_1^2 + x_2^4$  auf  $\mathbb{R}^2$  und den Punkt  $x^* = (0, 0)^T$ , dann gelten

$$\nabla f(x^*) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \nabla^2 f(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Da Satz 2.6 Ableitungen bis zur zweiten Ordnung verwendet, gibt er eine hinreichende Bedingung zweiter Ordnung an. ◇



**Fig. 3:** Ein striktes lokales Minimum mit semidefiniter Hessematrix.

**Satz 2.8. (Hinreichendes Kriterium zweiter Ordnung für einen Sattelpunkt)**

Seien  $X \subseteq \mathbb{R}^n$  offen und  $f : X \rightarrow \mathbb{R}$  zweimal stetig differenzierbar.

Für  $x^* \in X$  gelte  $\nabla f(x^*) = 0$  und  $\nabla^2 f(x^*)$  ist indefinit, so hat  $f$  in  $x$  kein lokales Optimum.  $x^*$  heißt **Sattelpunkt** von  $f$  auf  $X$ .

## 2.2. Konvexe Funktionen

**Definition 2.9.**

Eine Menge  $X \subseteq \mathbb{R}^n$  heißt **konvex**, wenn für alle  $x, y \in X$  und  $\lambda \in (0, 1)$  auch  $\lambda x + (1 - \lambda)y$  oder äquivalent  $y + \lambda(x - y)$  in  $X$  liegt.

Seien  $X \subseteq \mathbb{R}^n$  eine konvexe Menge und  $f : X \rightarrow \mathbb{R}$  eine Funktion.

1.  $f$  heißt **strikt konvex** bzw. **konvex**, wenn für alle  $x, y \in X$ ,  $x \neq y$  und alle  $\lambda \in (0, 1)$  gilt

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y) \quad \text{bzw.} \quad f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

2.  $f$  heißt **gleichmäßig konvex** mit **Modul**  $\mu > 0$ , falls für alle  $x, y \in X$ ,  $\lambda \in (0, 1)$  gilt:

$$f(\lambda x + (1 - \lambda)y) + \mu \lambda(1 - \lambda) \|x - y\|^2 \leq \lambda f(x) + (1 - \lambda)f(y).$$

**Bemerkung 2.10.**

Aus der Definition folgt, dass jede gleichmäßig konvexe Funktion auch strikt konvex ist und jede strikt konvexe Funktion konvex ist. Die Umkehrungen gelten im Allgemeinen nicht.  $\diamond$

Ist  $f$  eine differenzierbare Funktion, so können wir Konvexität wie folgt charakterisieren; vergleiche dazu [8, Satz 6.3].

**Satz 2.11.**

Sei  $f : X \rightarrow \mathbb{R}$  stetig differenzierbar und  $X \subseteq \mathbb{R}^n$  offen und konvex. Dann gelten:

1.  $f$  ist genau dann konvex, wenn für alle  $x, y \in X$  mit  $x \neq y$  gilt:

$$\nabla f(x)^T(y - x) \leq f(y) - f(x) \quad (2.1)$$

2.  $f$  ist strikt konvex genau dann, wenn für alle  $x, y \in X$  mit  $x \neq y$

$$\nabla f(x)^T(y - x) < f(y) - f(x)$$

erfüllt ist.

3. Die Funktion  $f$  ist gleichmäßig konvex genau dann, wenn es  $\mu > 0$  gibt mit

$$\nabla f(x)^T(y - x) + \mu\|y - x\|^2 \leq f(y) - f(x) \quad \text{für alle } x, y \in X$$

**Beweis.**

1.

„ $\Rightarrow$ “ Sei  $f$  konvex für beliebige  $x, y \in X$  und für alle  $\lambda \in (0, 1]$  bekommen wir

$$\frac{1}{\lambda}(f(x + \lambda(y - x)) - f(x)) \leq \frac{1}{\lambda}((1 - \lambda)f(x) + \lambda f(y) - f(x)) = f(y) - f(x)$$

Bilden wir den Grenzwert  $\lambda \rightarrow 0^+$ , so gilt

$$\nabla f(x)^T(y - x) = \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda}(f(x + \lambda(y - x)) - f(x)) \leq f(y) - f(x) \quad \text{also} \quad (2.1)$$

„ $\Leftarrow$ “ Sei (2.1) erfüllt. Wir wählen  $x, y \in X$  und  $\lambda \in [0, 1]$  beliebig und definieren  $x_\lambda = (1 - \lambda)x + \lambda y$ . Um  $(1 - \lambda)f(x) + \lambda f(y) - f(x_\lambda) \geq 0$  zu zeigen, betrachten wir mit (2.1)

$$\begin{aligned} (1 - \lambda)f(x) + \lambda f(y) - f(x_\lambda) &= (1 - \lambda)(f(x) - f(x_\lambda)) + \lambda(f(y) - f(x_\lambda)) \\ &\stackrel{(2.1)}{\geq} (1 - \lambda)\nabla f(x_\lambda)^T(x - x_\lambda) + \lambda\nabla f(x_\lambda)^T(y - x_\lambda) \\ &= \nabla f(x_\lambda)^T((1 - \lambda)x + \lambda y - x_\lambda) \\ &= 0 \end{aligned} \quad (2.2)$$

was zu zeigen war.

2.

„ $\Rightarrow$ “ Sei  $f$  streng konvex. Wir wählen  $x, y \in X$  mit  $x \neq y$  und setzen  $z := \frac{1}{2}(x + y)$ . Dann erhalten wir  $f(z) < \frac{1}{2}(f(x) + f(y))$  also  $f(z) - f(x) < \frac{1}{2}(f(y) - f(x))$ . Nun verwenden wir 1. Es gilt daher

$$\begin{aligned} \nabla f(x)^T(y - x) &= 2\nabla f(x)^T(z - x) \\ &\leq 2(f(z) - f(x)) \\ &< f(y) - f(x) \end{aligned}$$

was zu zeigen war.

„ $\Leftarrow$ “ Diese Aussage bekommen wir sofort aus der Abschätzung (2.2), wenn wir  $>$  statt  $\geq$  verwenden.

3.

„ $\Rightarrow$ “ Wir betrachten wie in Teil 1. den Differenzenquotienten

$$\begin{aligned}\nabla f(x)^T(y-x) &= \lim_{\lambda \rightarrow 0^+} \frac{f(x + \lambda(y-x)) - f(x)}{\lambda} \\ &\leq \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} ((1-\lambda)f(x) + \lambda f(y) - \mu\lambda(1-\lambda)\|y-x\|^2 - f(x)) \\ &= f(y) - f(x) - \mu\|y-x\|^2\end{aligned}$$

„ $\Leftarrow$ “ Es gelten

$$\|x - x_\lambda\| = \|x - (1-\lambda)x - \lambda y\| = \lambda\|x - y\| \quad \text{und} \quad \|y - x_\lambda\| = \|y - (1-\lambda)x - \lambda y\| = (1-\lambda)\|x - y\|$$

Daher erhalten wir (vgl. das Vorgehen bei der Abschätzung (2.1))

$$\begin{aligned}(1-\lambda)f(x) + f(y) - f(x_\lambda) &= (1-\lambda)(f(x) - f(x_\lambda)) + \lambda(f(y) - f(x_\lambda)) \\ &\geq (1-\lambda)(\nabla f(x_\lambda)^T(x - x_\lambda) + \mu\|x - x_\lambda\|^2) + \lambda(\nabla f(x_\lambda)^T(y - x_\lambda) \\ &\quad + \mu\|y - x_\lambda\|^2) \\ &= \nabla f(x_\lambda)^T((1-\lambda)x + \lambda y - x_\lambda) + \mu((1-\lambda)\|x - x_\lambda\|^2 + \lambda\|y - x_\lambda\|^2) \\ &= \mu((1-\lambda)\lambda^2 + \lambda(1-\lambda)^2)\|x - y\|^2 \\ &= \mu(1-\lambda)\lambda\|y - x\|^2\end{aligned}$$

was zu zeigen war. □

Ist das Zielfunktional  $f$  zweimal stetig differenzierbar, so können wir die Konvexität von  $f$  mittels der Hessematrix von  $f$  charakterisieren. Ohne Beweis geben wir dazu folgende Charakterisierung zweimal stetig differenzierbarer Funktionen an; vergleiche dazu [8, Satz 6.4].

**Satz 2.12. (Charakterisierung konvexer Funktionen)**

Seien  $X \subseteq \mathbb{R}^n$  offen und konvex sowie  $f : X \rightarrow \mathbb{R}$  zweimal stetig differenzierbar. Dann gelten:

1.  $f$  ist konvex auf  $X \Leftrightarrow \nabla^2 f(x)$  ist für alle  $x \in X$  positiv semidefinit.
2. Ist  $\nabla^2 f(x)$  für alle  $x \in X$  positiv definit, so ist  $f$  strikt konvex auf  $X$ .
3.  $f$  ist genau dann gleichmäßig konvex auf  $X$ , wenn  $\nabla^2 f(x)$  gleichmäßig positiv definit ist, d.h. wenn es ein  $\mu > 0$  gibt, so dass für alle  $x \in X$  und alle  $d \in \mathbb{R}^n$  gilt:  $d^T \nabla^2 f(x) d \geq \mu \|d\|^2$ .

**Beweis.**

1.

„ $\Rightarrow$ “ Sei also  $f$  konvex und  $x \in X$ ,  $d \in \mathbb{R}^n$  beliebig. Da  $X$  offen ist, gibt es  $\tau = \tau(x, d) > 0$  mit  $x + td \in X$  für alle  $t \in [0, \tau]$ . Für  $0 < t \leq \tau$  folgt aus Satz 2.11 mit (2.1) und TAYLOREntwicklung

$$0 \leq f(x + td) - f(x) - t\nabla f(x)^T d = \frac{t^2}{2} d^T \nabla^2 f(x) d + o(t^2)$$

Wir erhalten die Behauptung nach Multiplikation mit  $\frac{2}{t^2}$  und dem Grenzübergang  $t \rightarrow 0^+$ .

„ $\Leftarrow$ “ Seien  $x, y \in X$  beliebig. Mit TAYLOREntwicklung folgt für ein  $\vartheta \in [0, 1]$

$$\begin{aligned} f(y) - f(x) &= \nabla f(x)^T(y - x) + \frac{1}{2}(y - x)^T \nabla^2 f(x + \vartheta(y - x))(y - x) \\ &\geq \nabla f(x)^T(y - x). \end{aligned} \quad (2.3)$$

Aus Satz 2.11 mit (2.1) folgt, dass  $f$  konvex ist.

2. Wir erhalten die Aussagen aus (2.3), wenn wir  $x, y \in X$  mit  $x \neq y$  beliebig wählen und  $\geq$  durch  $>$  ersetzen.

3.

„ $\Rightarrow$ “ Wie in Teil 1. erhalten wir für beliebige  $x \in X$  und  $d \in \mathbb{R}^n \setminus \{0\}$  ein  $\tau = \tau(x, d) > 0$ , so dass

$$0 \leq f(x + td) - f(x) - t\nabla f(x)^T d - \mu \|td\|^2 = \frac{t^2}{2} d^T \nabla^2 f(x) d - t^2 \mu \|d\|^2 + \mathcal{O}(t^2)$$

für alle  $0 < t \leq \tau$ . Nach Multiplikation mit  $\frac{2}{t^2}$  und Grenzübergang  $t \rightarrow 0^+$  folgt

$$d^T \nabla^2 f(x) d \geq \mu \|d\|^2.$$

„ $\Leftarrow$ “ Seien  $x \in X$ ,  $d \in \mathbb{R}^n$  beliebig gewählt. Nach Satz von TAYLOR existiert ein  $\vartheta \in [0, 1]$  mit

$$\begin{aligned} f(y) - f(x) &= \nabla f(x)^T(y - x) + \frac{1}{2}(y - x)^T \nabla^2 f(x + \vartheta(y - x))(y - x) \\ &\geq \nabla f(x)^T(y - x) + \frac{\mu}{2} \|y - x\|^2 \end{aligned}$$

da  $\nabla^2 f$  gleichmäßig positiv definit ist. Daraus folgt, dass  $f$  gleichmäßig konvex ist. □

### Bemerkung 2.13.

- Die 2. Aussage von Satz 2.12 lässt sich nicht ohne Weiteres umkehren: Die Funktion  $f(x) = x^4$  mit  $X = \mathbb{R}$  beispielsweise ist strikt konvex, aber es gilt  $f''(0) = 0$ .
- $f: \mathbb{R}^n \rightarrow \mathbb{R}$  sei eine quadratische Funktion, d.h. von der Gestalt

$$f(x) = \frac{1}{2} x^T Q x + c^T x + \gamma$$

mit symmetrischer Matrix  $Q \in \mathcal{S}_n$ , Vektor  $c \in \mathbb{R}^n$  und Skalar  $\gamma \in \mathbb{R}$ . Dann gelten:

- $f$  ist genau dann konvex, wenn  $Q$  positiv semidefinit ist.
- $f$  ist strikt konvex  $\Leftrightarrow f$  ist gleichmäßig konvex  $\Leftrightarrow Q$  ist positiv definit. ◇

Wir kommen nun zum Optimierungsproblem zurück; vergleiche dazu [8, Satz 6.5]

#### Satz 2.14. (Existenz und Eindeutigkeit für konvexe Optimierungsaufgaben)

Seien  $f: X \rightarrow \mathbb{R}$  konvex, stetig differenzierbar und  $X \subseteq \mathbb{R}^n$  konvex. Wir betrachten das Optimierungsproblem

$$\min f(x) \quad \text{u.d.N.} \quad x \in X \quad (2.4)$$

Dann gelten:

- Jedes lokale Minimum von  $f$  ist ein globales Minimum.
- Die Lösungsmenge von (2.4) ist konvex (gegebenenfalls leer).
- Ist  $f$  strikt konvex auf  $X$ , so besitzt (2.4) höchstens eine Lösung. Diese ist - im Falle der Existenz - nach Teil 1. das strikte globale Minimum von  $f$  auf  $X$ .
- Ist  $X$  offen und  $x^* \in X$  ein stationärer Punkt von  $f$ , so ist  $x^*$  ein globales Minimum von  $f$  auf  $X$ .

**Beweis.**

1. Sei  $x^*$  ein lokales Minimum von  $f$  auf  $X$ . Angenommen, es existiert ein  $x \in X$  mit  $f(x) < f(x^*)$ . Dann folgt für alle  $\lambda \in [0, 1]$

$$\begin{aligned} f(x^* + \lambda(x - x^*)) &\leq (1 - \lambda)f(x^*) + \lambda f(x) \\ &< (1 - \lambda)f(x^*) + \lambda f(x^*) = f(x^*) \end{aligned}$$

d.h., aber  $x^*$  ist kein lokales Minimum, was ein Widerspruch ist.

2. Seien  $x^1, x^2$  Lösungen von (2.4), also  $f(x^1) = f(x^2) = \min_{x \in X} f(x)$ . Für  $\lambda \in (0, 1)$  gelten dann

$$\lambda x^1 + (1 - \lambda)x^2 \in X \quad \text{und} \quad f(\lambda x^1 + (1 - \lambda)x^2) \leq \lambda f(x^1) + (1 - \lambda)f(x^2) = \min_{x \in X} f(x).$$

Also nimmt  $f$  auch an  $\lambda x^1 + (1 - \lambda)x^2$  sein Minimum an.

3. Angenommen, (2.4) hat zwei verschiedene Lösungen  $x^1, x^2$  mit  $f(x^1) = f(x^2) = \min_{x \in X} f(x)$ . Für  $\lambda \in (0, 1)$  gelten dann

$$\lambda x^1 + (1 - \lambda)x^2 \in X \quad \text{und} \quad f(\lambda x^1 + (1 - \lambda)x^2) < \lambda f(x^1) + (1 - \lambda)f(x^2) = \min_{x \in X} f(x),$$

was einen Widerspruch ergibt.

4. Für alle  $x \in X$  folgt aus Satz 2.11 mit (2.1)

$$f(x) - f(x^*) \geq \nabla f(x^*)^T (x - x^*) = 0$$

Somit ist  $x^*$  ein globales Minimum von  $f$  auf  $X$

□

**Bemerkung 2.15.**

- Das Problem (2.4) muss selbst für strikt konvexes  $f$  keine Lösung besitzen. Betrachte dazu zum Beispiel  $f(x) = \exp(x)$  auf  $X = \mathbb{R}$ .
- Aus der 4. Aussage von Satz 2.14 folgt, dass  $\nabla f(x^*) = 0$  auch hinreichend dafür ist, dass  $x^*$  eine globale Minimalstelle ist.

◇

**3. Abstiegsverfahren und Schrittweitenstrategien**

Wir betrachten ein Abstiegsverfahren zur Lösung des unrestringierten Optimierungsproblems

$$\min_{x \in \mathbb{R}^n} f(x)$$

mit stetig differenzierbarer Funktion  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Die zentrale Idee der Verfahren in diesem Abschnitt ist wie folgt:

- Ist man in einem Punkt  $x \in \mathbb{R}^n$ , so sucht man eine Richtung  $d \in \mathbb{R}^n$  aus, längs welcher der Funktionswert fällt (**Abstiegsverfahren**).
- Entlang dieser Richtung  $d$  geht man so lange, bis man den Funktionswert von  $f$  hinreichend verkleinert hat (**Schrittweitenstrategie**).

Diese Schritte wollen wir formalisieren.

**Definition 3.1.**

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  und  $x \in \mathbb{R}^n$ . Ein Vektor  $d \in \mathbb{R}^n$  heißt **Abstiegsrichtung** von  $f$  im Punkt  $x$ , wenn es ein  $t_0 > 0$  gibt mit

$$f(x + td) < f(x) \quad \text{für alle} \quad t \in (0, t_0].$$

**Bemerkung 3.2.** (Hinreichende Abstiegsbedingung)

Ist  $f$  stetig differenzierbar, dann ist die Bedingung

$$\nabla f(x)^T d < 0 \quad (3.1)$$

hinreichend dafür, dass  $d \in \mathbb{R}^n$  eine Abstiegsrichtung von  $f$  in  $x$  ist.

Um dies einzusehen, definieren wir  $\varphi(t) = f(x + td)$ . Aus  $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$  folgt dann

$$\varphi(t) = \varphi(0) + t\varphi'(0) + r(t) \quad \text{mit} \quad \frac{1}{t}r(t) \rightarrow 0 \quad \text{für} \quad t \searrow 0,$$

d.h.  $r(t) = \mathcal{O}(t)$  in der Landau-Symbol-Formalisierung. Es gelten  $\varphi(0) = f(x)$  und  $\varphi'(0) = \nabla f(x)^T d < 0$ . Also existiert ein  $t_0 > 0$  mit

$$\frac{\varphi(t) - \varphi(0)}{t} = \nabla f(x)^T d + \frac{r(t)}{t} \quad \implies \quad \forall t \in (0, t_0]: \quad \frac{\varphi(t) - \varphi(0)}{t} < 0,$$

d.h.  $f(x + td) < f(x)$  und somit ist  $d$  eine Abstiegsrichtung von  $f$  in  $x$ .  $\diamond$

**Bemerkung 3.3.**

1. Die Bedingung (3.1) bedeutet, dass der Winkel  $\varphi$  zwischen  $d$  und dem negativen Gradienten  $-\nabla f(x)$  von  $f$  in  $x$  weniger als  $\frac{\pi}{2}$  beziehungsweise  $90^\circ$  beträgt. Betrachte dazu:

$$0 > \nabla f(x)^T d \quad \implies \quad 0 < -\nabla f(x)^T d = \cos(\varphi) \underbrace{\|-\nabla f(x)\| \|d\|}_{>0}$$

Es folgt  $\cos(\varphi) > 0$  und damit  $\varphi \in [0, \frac{\pi}{2})$ .

2. Das Kriterium (3.1) ist nicht notwendig. Ist  $x$  beispielsweise eine strikte lokale Maximalstelle, so sind alle  $d \in \mathbb{R}^n$  Abstiegsrichtungen von  $f$  in  $x$ , aber (3.1) ist nicht für alle diese Richtungen erfüllt.  $\diamond$

**Beispiel 3.4.**

1.  $d = -\nabla f(x)$  ist stets eine Abstiegsrichtung, sofern  $x$  kein stationärer Punkt ist, denn dann gilt  $\nabla f(x)^T d = -\|\nabla f(x)\|^2 < 0$ .  $-\nabla f(x)$  ist sogar die Richtung des steilsten Abstiegs: Unter allen Richtungen  $d \in \mathbb{R}^n$  mit  $\|d\| = 1$  ist der normierte Gradient von  $f$  in  $x$  Minimalstelle von  $\min \nabla f(x)^T d$  u.d.N.  $\|d\| = 1$ . Wählt man  $d = -\nabla f(x)$ , so bezeichnet man das Abstiegsverfahren als **Gradientenverfahren**.
2. Ist  $M \in \mathcal{S}_n$  positiv definit, dann ist  $d = -M\nabla f(x)$  im Fall  $\nabla f(x) \neq 0$  eine Abstiegsrichtung, denn es gilt  $\nabla f(x)^T d = -\nabla f(x)^T M\nabla f(x) < 0$ . Man bezeichnet die Abstiegsverfahren dann als **gradientenähnliche Verfahren**.  $\diamond$

**3.1. Abstiegsverfahren**

Wir wollen in algorithmischer Form ein allgemeines Abstiegsverfahren angeben. Hier werden iterativ Punkte  $x^k \in \mathbb{R}^n$  erzeugt mit  $f(x^{k+1}) < f(x^k)$ , bis eine Abbruchbedingung beispielsweise von der Form  $\|x^k - x^*\| < \epsilon$  für ein  $\epsilon > 0$  erfüllt ist.

In theoretischen Konvergenzuntersuchungen betrachten wir kein Konvergenzkriterium. Vielmehr nehmen wir an, dass eine unendliche Folge  $(x^k)_{k \in \mathbb{N}}$  erzeugt wird, und untersuchen diese Folge auf Grenzwerte bzw. Häufungspunkte.

**Satz 3.6.** (Hinreichendes Konvergenzkriterium des Allgemeinen Abstiegsverfahrens)

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar und  $(x^k)_{k \in \mathbb{N}}$  eine durch Algorithmus 3.5 erzeugte Folge. Weiter mögen gelten:

**Algorithmus 3.5** (Allgemeines Abstiegsverfahren)**Eingabe:** Startpunkt  $x_0 \in \mathbb{R}^n$ .

- 1: Setze  $k = 0$ .
- 2: **while** Konvergenzkriterium nicht erfüllt **do**
- 3: Bestimme Abstiegsrichtung  $d^k$  von  $f$  in  $x^k$ .
- 4: Bestimme eine Schrittweite  $t_k > 0$  mit  $f(x^k + t_k d^k) < f(x^k)$ .
- 5: Setze  $x^{k+1} = x^k + t_k d^k$  und  $k = k + 1$ .
- 6: **end while**

1. Es gibt eine Konstante  $\theta_1 > 0$ , so dass die folgende **Winkelbedingung** für alle  $k \in \mathbb{N}$  erfüllt ist:

$$-\nabla f(x^k)^\top d^k \geq \theta_1 \|\nabla f(x^k)\| \|d^k\|$$

2. Es gibt eine von  $(x^k)_{k \in \mathbb{N}}$  und  $(d^k)_{k \in \mathbb{N}}$  unabhängige Konstante  $\theta_2 > 0$  mit der folgenden Bedingung des **hinreichenden Abstiegs**:

$$f(x^k + t_k d^k) \leq f(x^k) - \theta_2 \left( \frac{\nabla f(x^k)^\top d^k}{\|d^k\|} \right)^2 \quad \text{mit } t_k > 0 \text{ für alle } k \in \mathbb{N}.$$

Dann ist jeder Häufungspunkt der Folge  $(x^k)_{k \in \mathbb{N}}$  ein stationärer Punkt von  $f$ .

**Beweis.**

Sei  $x^*$  ein Häufungspunkt der durch Algorithmus 3.5 erzeugten Folge  $(x^k)_{k \in \mathbb{N}}$ . Dann gibt es eine Teilfolge  $(x^{k_\nu})_{\nu \in \mathbb{N}}$  mit  $x^{k_\nu} \rightarrow x^*$  für  $\nu \rightarrow \infty$ . Nach Konstruktion ist  $(f(x^k))_{k \in \mathbb{N}}$  monoton fallend, also folgt aus  $f(x^{k_\nu}) \rightarrow f(x^*)$  für  $\nu \rightarrow \infty$ , dass auch die gesamte Folge der Funktionswerte  $(f(x^k))_{k \in \mathbb{N}}$  gegen  $f(x^*)$  konvergiert.

Weiter liefern die beiden Bedingungen des Satzes die Abschätzung

$$0 \stackrel{k \rightarrow \infty}{\leftarrow} f(x^{k+1}) - f(x^k) \leq -\theta_2 \left( \frac{\nabla f(x^k)^\top d^k}{\|d^k\|} \right)^2 \leq -\theta_1^2 \theta_2 \|\nabla f(x^k)\|^2 \leq 0,$$

also ist  $\lim_{k \rightarrow \infty} \|\nabla f(x^k)\| = 0$ . Damit ist jeder Häufungspunkt von  $f$  auch ein stationärer Punkt von  $f$ .  $\square$

**Bemerkung 3.7.**

Bezeichne  $\eta_k$  den Winkel zwischen  $d^k$  und  $-\nabla f(x^k)$ , dann bedeutet die Winkelbedingung aus Satz 3.6, dass

$$\cos(\eta_k) = \frac{-\nabla f(x^k)^\top d^k}{\|\nabla f(x^k)\| \|d^k\|}$$

gleichmäßig größer als 0 ist. Dies impliziert, dass  $d^k$  und  $-\nabla f(x^k)$  nur in einem Winkel zwischen 0 und weniger als  $\frac{\pi}{2}$  zueinander stehen. Ein wichtiges Beispiel einer Abstiegsrichtung, für welche diese Bedingung erfüllt ist, ist die Wahl  $d^k = -\nabla f(x^k)$ .  $\diamond$

**3.2. Schrittweitenstrategien**

Das Allgemeine Abstiegsverfahren (Algorithmus 3.5) besitzt in der Wahl der Abstiegsrichtung  $d^k$  und der Schrittweite  $t_k > 0$  große Freiheitsgrade. Die nahe liegende Minimierungsregel  $t_k = t_k^{\min}$  für die zu wählende Schrittweite mit

$$f(x^k + t_k^{\min} d^k) = \min_{t \geq 0} f(x^k + t d^k)$$

ist unter der Annahme, dass die Niveaumenge  $\mathcal{L}(x^0) = \{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\}$  kompakt ist und  $\nabla f(x)$  Lipschitz-stetig auf  $\mathcal{L}(x^0)$  ist, wohldefiniert. Allerdings ist diese Regel im Allgemeinen nicht praktikabel: In jedem Schritt müsste dazu schließlich ein eindimensionales Optimierungsproblem exakt gelöst werden.

Wir betrachten im Folgenden **Schrittweitenstrategien**, für welche die zugehörige Folge  $(x^k)_{k \in \mathbb{N}}$  gegen  $x^*$  konvergiert.

### 3.2.1. Die Armijo-Regel

Sei ein  $\alpha \in [0, 1]$  fest vorgegeben. Die **Armijo-Regel** ist eine Bedingung, die einen hinreichenden Abstieg sichert, und lautet

$$f(x + td) \leq f(x) + \alpha t \nabla f(x)^T d. \quad (3.2)$$

Wegen der stetigen Differenzierbarkeit von  $f$  und der Eigenschaft  $\nabla f(x)^T d < 0$  ist die Existenz eines  $t_0 > 0$  gesichert, so dass für alle  $t \in (0, t_0]$  die Bedingung (3.2) erfüllt ist; siehe [8, Lemma 7.5].

#### Lemma 3.8.

Seien  $f : U \rightarrow \mathbb{R}$  stetig differenzierbar auf der offenen Menge  $U \subset \mathbb{R}^n$ ,  $x \in U$ ,  $\alpha \in (0, 1)$  und  $d \in \mathbb{R}^n$  eine Richtung mit  $\nabla f(x)^T d < 0$ . Dann existiert ein  $\bar{t} > 0$  mit

$$f(x + td) \leq f(x) + \alpha t \nabla f(x)^T d \quad \text{für alle } t \in [0, \bar{t}] \quad (3.3)$$

#### Beweis.

Für  $t = 0$  ist (3.3) offensichtlich erfüllt. Sei nun  $t > 0$  hinreichend klein, so dass  $x + td \in U$  gilt. Dann erhalten wir

$$\frac{1}{t}(f(x + td) - f(x)) - \alpha \nabla f(x)^T d \xrightarrow{t \rightarrow 0^+} \nabla f(x)^T d(1 - \alpha) < 0,$$

wobei wir die Voraussetzungen  $\nabla f(x)^T d < 0$  und  $\alpha \in (0, 1)$  genutzt haben. Wir können also  $\bar{t}$  so klein wählen, dass

$$\frac{1}{t}(f(x + td) - f(x)) - \alpha \nabla f(x)^T d \leq 0$$

für alle  $t \in (0, \bar{t}]$  gilt. Dieses  $\bar{t}$  erfüllt (3.3). Zusammenfassend gilt also (3.3) für alle  $t \in [0, \bar{t}]$ .  $\square$

Wir visualisieren die Armijo-Regel (3.2): Die durchgezogene Linie repräsentiert den Graphen der 1D-Abbildung  $f_1 : t \mapsto f(x + td)$  für  $t \geq 0$  während die gestrichelte Gerade die affin lineare Funktion  $f_2 : t \mapsto f(x) + \alpha t \nabla f(x)^T d$  darstellt.

Da  $d$  eine Abstiegsrichtung ist, ist die Existenz von  $b > 0$  gesichert, so dass  $f_1(t) \leq f_2(t)$  für alle  $t \in [a, b]$  erfüllt ist.

Zufällig ist die Armijo-Bedingung auch für alle  $t \in [c, d]$  erfüllt (dies wäre beispielsweise bei einer strikt konvexen Funktion nicht der Fall).

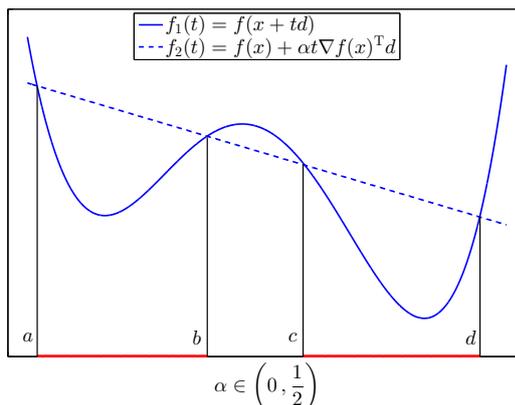


Fig. 4: Die Armijo-Regel.

Zur tatsächlichen Berechnung einer geeigneten **Schrittweite**  $t$  überprüft man (3.2) sequenziell zum Beispiel für

$$t = \beta^l, \quad l = 0, 1, 2, \dots \quad (3.4)$$

mit einem fest vorgegebenen  $\beta \in (0, 1)$ , etwa  $\beta = \frac{1}{2}$ . Bei erstmaliger Gültigkeit von (3.2) bricht man ab. Im Folgenden verallgemeinern wir die Wahl der Schrittweite: Ist (3.2) für ein  $t_c$  nicht erfüllt, dann wird ein  $t_+ < t_c$  so konstruiert, dass  $t_+ \in [\nu_0 t_c, \nu^0 t_c]$  gilt für gewisse Parameter  $0 < \nu_0 \leq \nu^0 < 1$ .

**Algorithmus 3.9 (Armijo-Schrittweitenalgorithmus)**

**Eingabe:** Abstiegsrichtung  $d$ , Punkt  $x$ , Parameter  $\alpha \in (0, 1)$ ,  $0 < \nu_0 \leq \nu^0 < 1$ .

- 1: Setze  $l = 0$  und  $t^{(0)} = 1$ .
- 2: **while** Armijo-Bedingung (3.2) nicht erfüllt **do**
- 3:   Bestimme  $t^{(l+1)} \in [\nu_0 t^{(l)}, \nu^0 t^{(l)}]$  und setze  $l = l + 1$ .
- 4: **end while**

Im Weiteren behandeln wir **gradientenähnliche Richtungen** d.h.  $d = -M\nabla f(x)$  mit positiv definitem  $M \in \mathcal{S}_n$  für die Armijo Schrittweitenstrategie. Bei gradientenähnlichen Verfahren lässt sich die Anzahl der Iterationen von Algorithmus 3.9 abschätzen.

**Lemma 3.10. (Endliche Terminierungsbedingung des Armijo-Schrittweitenalgorithmus)**

Seien  $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ ,  $\nabla f$  Lipschitz-stetig zur Konstante  $L$ ,  $\alpha \in (0, 1)$ ,  $x \in \mathbb{R}^n$  mit  $\nabla f(x) \neq 0$  und  $M \in \mathcal{S}_n$  positiv definit. Bezeichne  $\lambda_s$  den kleinsten und  $\lambda_g$  den größten Eigenwert von  $M^{-1}$  und  $\kappa_2(M^{-1}) = \|M^{-1}\|_2 \|M\|_2 = \frac{\lambda_g}{\lambda_s}$  die spektrale Konditionszahl von  $M^{-1}$ . Dann ist (3.2) für alle  $0 < t \leq 2\lambda_s(1 - \alpha)(L\kappa_2(M^{-1}))^{-1}$  erfüllt.

**Beweis.**

Es gilt  $f(x + td) - f(x) = t \int_0^1 \nabla f(x + \tau td)^T d d\tau$ . Wir erhalten daraus

$$f(x + td) = f(x) + t \nabla f(x)^T d + t \int_0^1 (\nabla f(x + \tau td) - \nabla f(x))^T d d\tau.$$

Aus der Lipschitz-Stetigkeit des Gradienten von  $f$  folgt

$$f(x + td) \leq f(x) + t \nabla f(x)^T d + \frac{Lt^2}{2} \|d\|^2.$$

Wegen  $z^T M^2 z \leq \lambda_s^{-2} \|z\|^2$  und  $\|z\|^2 \leq \lambda_g z^T M z$  für alle  $z \in \mathbb{R}^n$  können wir  $\|d\|^2$  abschätzen via

$$\begin{aligned} \|d\|^2 &= \|M \nabla f(x)\|^2 = \nabla f(x)^T M^2 \nabla f(x) \leq \lambda_s^{-2} \|\nabla f(x)\|^2 \\ &\leq \lambda_g \lambda_s^{-2} \nabla f(x)^T M \nabla f(x) = -\kappa_2(M^{-1}) \lambda_s^{-1} \nabla f(x)^T d. \end{aligned}$$

Einsetzen ergibt

$$f(x + td) \leq f(x) + t \left( 1 - \frac{\kappa_2(M^{-1}) Lt}{\lambda_s} \right) \nabla f(x)^T d.$$

Die Armijo-Bedingung (3.2) ist also erfüllt, falls  $\alpha \leq (1 - \frac{\kappa_2(M^{-1}) Lt}{\lambda_s})/2$  bzw.  $t \leq 2\lambda_s(1 - \alpha)(L\kappa_2(M^{-1}))^{-1}$  gilt.  $\square$

**Lemma 3.11. (Endliche Terminierung & maximale Iterationszahl für gradientenähnliche Richtungen)**

Seien  $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ ,  $\nabla f$  Lipschitz-stetig zur Konstante  $L$ ,  $\alpha \in (0, 1)$ ,  $(x^k)_{k \in \mathbb{N}} \subseteq \mathbb{R}^n$  eine durch das Allgemeine Abstiegsverfahren 3.5 mit Armijo-Schrittweitenwahl 3.9 generierte Iterationsfolge. Die Folge  $(M^k)_{k \in \mathbb{N}} \subseteq \mathcal{S}_n$  der zu den Abstiegsrichtungen  $d^k = -M^k \nabla f(x^k)$  gehörigen, positiv definiten Matrizen erfülle  $\lambda_0 \leq \lambda_s^k \leq \lambda_g^k \leq \lambda^0$  für geeignete  $0 < \lambda_0 \leq \lambda^0$ . Dann gelten:

1. Jede Schrittweite  $t_k$  genügt der Abschätzung  $t_k \geq \tilde{t} = 2\nu_0 \lambda_0 (1 - \alpha) / (L\tilde{\kappa}_2)$  mit  $\tilde{\kappa}_2 = \lambda^0 / \lambda_0$ .
2. Jede Schrittweitensuche terminiert nach höchstens  $\log(2\lambda_0(1 - \alpha) / (L\tilde{\kappa}_2)) / \log \nu^0$  Iterationen.

**Beweis.**

1. Mit  $\lambda_s^k \geq \lambda_0 > 0$  und  $\kappa_2^k = \kappa_2((M^k)^{-1}) = \lambda_g^k / \lambda_s^k \leq \lambda^0 / \lambda_0 = \tilde{\kappa}_2$  folgt aus Lemma 3.11:

$$\forall k \in \mathbb{N} : \frac{2\lambda_s^k(1 - \alpha)}{L\kappa_2^k} \geq \frac{2\lambda_0(1 - \alpha)}{L\tilde{\kappa}_2} > 0 \quad \implies \quad \forall k \in \mathbb{N} : t_k \geq \tilde{t}.$$

2. Wegen  $t^{(0)} = 1$  und  $t^{(k+1)} \leq \nu^0 t^{(k)}$  wird eine zulässige Schrittweite  $t^{(k)}$  spätestens nach  $m$  Schritten gefunden mit

$$t^{(m)} \leq \nu^0 t^{(m-1)} \leq (\nu^0)^m t^{(0)} = (\nu^0)^m \leq \frac{2\lambda_0(1-\alpha)}{L\tilde{\kappa}_2}$$

$$\implies m = \log_{\nu^0} \left( \frac{2\lambda_0(1-\alpha)}{L\tilde{\kappa}_2} \right) = \frac{1}{\log \nu^0} \log \left( \frac{2\lambda_0(1-\alpha)}{L\tilde{\kappa}_2} \right).$$

□

**Satz 3.12. (Hinreichendes Konvergenzkriterium des Abstiegsverfahrens mit Armijo-Schrittweiten)**

Seien  $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ ,  $\nabla f$  Lipschitz-stetig zur Konstante  $L$ ,  $\alpha \in (0, 1)$ ,  $(x^k)_{k \in \mathbb{N}} \subseteq \mathbb{R}^n$  eine durch das Allgemeine Abstiegsverfahren 3.5 mit Armijo-Schrittweitenwahl 3.9 generierte Iterationsfolge. Die Folge  $(M^k)_{k \in \mathbb{N}} \subseteq \mathcal{S}_n$  der zu den Abstiegsrichtungen  $d^k = -M^k \nabla f(x^k)$  gehörigen, positiv definiten Matrizen erfülle  $\lambda_0 \leq \lambda_s^k \leq \lambda_g^k \leq \lambda^0$  für geeignete  $0 < \lambda_0 \leq \lambda^0$ .

Ist  $(f(x^k))_{k \in \mathbb{N}}$  von unten beschränkt, so ist jeder Häufungspunkt von  $f$  ein stationärer Punkt von  $f$ .

**Beweis.**

Sei  $(f(x^k))_{k \in \mathbb{N}}$  von unten beschränkt. Da die Folge nach Konstruktion zugleich monoton fällt, folgt die Konvergenz und mit Lemma 3.11

$$0 \leftarrow f(x^{k+1}) - f(x^k) < -\alpha t_k \nabla f(x^k)^\top M^k \nabla f(x^k) \leq -\alpha \tilde{t} \lambda_0^{-1} \|\nabla f(x^k)\|^2 \leq -\alpha \frac{2\nu_0(1-\alpha)}{L\tilde{\kappa}_2(M^{-1})} \|\nabla f(x^k)\|^2 \leq 0,$$

also ist  $\lim_{k \rightarrow \infty} \|\nabla f(x^k)\| = 0$ . □

**Bemerkung 3.13.**

1. Im Allgemeinen gibt es keine Garantie, dass ein (eindeutiger) Häufungspunkt existiert.
2. Die Variante (3.4) wird *Backtracking*-Strategie genannt.
3. Weitere Strategien basieren auf Polynommodellen, die  $\varphi(t) = f(x + td)$  durch ein *quadratisches* oder *kubisches Modell* ersetzen und dann dieses Modell minimieren.

◇

**3.2.2. Die WOLFE-POWELL-Regel**

Seien  $\alpha \in (0, \frac{1}{2})$  und  $\varrho \in [\alpha, 1]$  gegeben. Die **WOLFE-POWELL-Regel** lautet:

Zum Punkt  $x \in \mathbb{R}^n$  und zum Richtungsvektor  $d \in \mathbb{R}^n$  mit  $\nabla f(x)^\top d < 0$  bestimme eine Schrittweite  $t > 0$  so, dass

$$f(x + td) \leq f(x) + \alpha t \nabla f(x)^\top d \tag{3.5a}$$

$$\nabla f(x + td)^\top d \geq \varrho \nabla f(x)^\top d \tag{3.5b}$$

gelten bzw. gleichbedeutend  $\varphi(t) \leq \varphi(0) + \alpha t \varphi'(0)$  und  $\varphi'(t) \geq \varrho \varphi'(0)$ .

Zusätzlich zur Armijo-Regel (3.5a), welche sicherstellt, dass die gewählte Schrittweite  $t$  *hinreichend klein* ist, wird also gefordert, dass der Schritt nicht *zu klein* wird (3.5b).

Diese Bedingung wird dadurch motiviert, dass an einem lokalen Minimum  $t^*$  von  $\varphi$  gilt

$$\varphi'(t^*) = \nabla f(x + td)^T d = 0$$

und es daher sinnvoll ist, zu verhindern, dass  $\varphi$  bei  $t$  zu steil fällt.

In der Grafik ist die Armijo-Bedingung auf ganz  $[a, b]$  erfüllt, die WOLFE-POWELL-Regel allerdings nur auf dem Teilintervall  $[\bar{a}, b]$ . Weiterhin genügen auch alle Schrittweiten  $t \in [\bar{c}, d]$  den WOLFE-POWELL-Bedingungen.

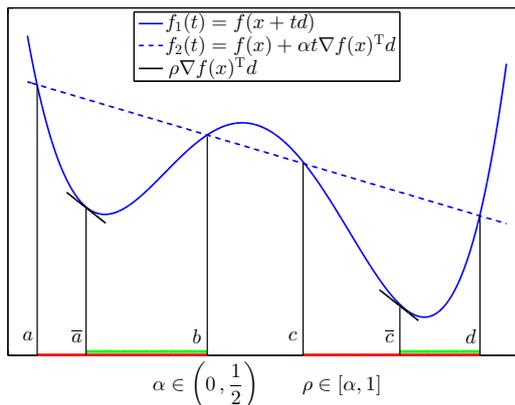


Fig. 5: Die Wolfe-Powell-Regel als modifizierte Armijo-Bedingung.

### Lemma 3.14.

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar,  $x \in \mathbb{R}^n$  und  $d \in \mathbb{R}^n$  mit  $\nabla f(x)^T d < 0$ . Ferner sei  $f$  an  $x$  in Richtung  $d$  beschränkt, d.h.,

$$\inf\{f(x + td) \mid t \geq 0\} > -\infty$$

Weiter seien  $\alpha \in (0, \frac{1}{2})$  und  $\rho \in (\alpha, 1)$ . Dann existiert ein  $t > 0$  mit (3.5a) und (3.5b).

### Beweis.

Die Funktion

$$\Psi(t) = f(x + td) - f(x) - \alpha t \nabla f(x)^T d$$

ist wegen  $\Psi'(0) = (1 - \alpha) \nabla f(x)^T d < 0$  in  $t = 0$  streng monoton fallend. Es gilt  $\Psi(0) = 0$  und somit gilt  $\Psi(t) < 0$  für kleine  $t > 0$  und die Menge  $\{t > 0 \mid \Psi(t) = 0\}$  ist abgeschlossen. Die Menge ist auch nichtleer, da  $f$  in  $x$  in Richtung  $d$  nach unten beschränkt ist. Daher existiert ein kleinstes  $\bar{t} > 0$  mit  $\Psi(\bar{t}) = 0$ . Wegen des Zwischenwertsatzes folgt  $\Psi(t) < 0$  für alle  $t \in (0, \bar{t})$ . Es gilt damit (3.5a) für  $t = \bar{t}$ . Ferner erhalten wir

$$\nabla f(x + \bar{t}d)^T d - \alpha \nabla f(x)^T d = \Psi'(\bar{t}) = \lim_{t \rightarrow 0^+} \frac{1}{t} \left( \underbrace{\Psi(\bar{t})}_{=0} - \underbrace{\Psi(\bar{t} - t)}_{<0} \right) \geq 0$$

Also gilt wegen  $0 < \alpha \leq \rho$

$$\nabla f(x + \bar{t}d)^T d \geq \alpha \nabla f(x)^T d \geq \rho \nabla f(x)^T d$$

Damit gilt auch (3.5b) für  $t = \bar{t}$ . □

Wir geben nun einen Algorithmus aus [8, Algorithmus S.3] an, der eine WOLFE-POWELL-Schrittweite berechnet. Die Idee des Verfahrens besteht darin, zuerst ein Intervall  $[t, \bar{t}]$  zu bestimmen, so dass (3.5a) für  $t = t$  gilt, (3.5a) aber für  $t = \bar{t}$  nicht gilt. Für

$$\Psi(t) = f(x + td) - f(x) - \alpha t \nabla f(x)^T d$$

gilt dann  $\Psi(t) \leq 0$  und  $\Psi(\bar{t}) \geq 0$ . Nach dem Zwischenwertsatz existiert also ein  $t^* \in [t, \bar{t}]$  mit  $\Psi(t^*) = 0$ . Durch Bisektion wird dieser Vorzeichenwechsel von  $\Psi$  eingeschachtelt, bis ein  $t \in (0, t^*]$  gefunden ist, welches (3.5) erfüllt. Offenbar gilt für  $t \leq t^*$  hinreichend klein (3.5).

Aus [8, Satz 9.4] zitieren wir den folgenden Konvergenzsatz mit Beweis.

**Algorithmus 3.15 (WOLFE-POWELL-Liniensuche)****Eingabe:**  $d \in \mathbb{R}^n$  mit  $\nabla f(x)^T d < 0$ 

- 1: Falls an  $t = 1$  die Armijo-Bedingung (3.5a) gilt, gehe zu Schritt 3.
- 2: Bestimme die größte Zahl  $\underline{t} \in \{2^{-1}, 2^{-2}, \dots\}$ , so dass für  $t = \underline{t}$  (3.5a) gilt. Setze  $\bar{t} = 2\underline{t}$  und gehe zu Schritt 5.
- 3: Falls  $t = 1$  (3.5b) erfüllt, **stop** mit Ergebnis  $t = 1$ . Sonst weiter mit Schritt 4.
- 4: Bestimme die kleinste Zahl  $\bar{t} \in \{2, 2^2, \dots\}$ , so dass (3.5a) verletzt ist für  $t = \bar{t}$ . Setze  $\underline{t} = \bar{t}/2$ .
- 5: Solange  $t = \underline{t}$  (3.5b) nicht erfüllt:
  - Berechne  $t = (\underline{t} + \bar{t})/2$ .
  - Erfüllt  $t$  (3.5a), setze  $\underline{t} = t$  sonst  $\bar{t} = t$ .
- 6: **stop** mit  $t = \underline{t}$

**Satz 3.16.**

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar,  $x \in \mathbb{R}^n$  und  $d \in \mathbb{R}^n$  mit  $\nabla f(x)^T d < 0$ . Ferner sei  $d$  eine Richtung an  $x$ , in die  $f$  nach unten beschränkt ist, d.h.,

$$\inf\{f(x + td) \mid t \geq 0\} > -\infty. \quad (3.6)$$

Weiter seien  $\alpha \in (0, \frac{1}{2})$  und  $\varrho \in (\alpha, 1)$  gegeben. Dann terminiert Algorithmus 3.15 nach endlich vielen Schritten mit einer Schrittweite  $t$ , die (3.5) erfüllt.

**Beweis.**

Schritt 2 implementiert die Armijo-Regel mit  $\beta = 1/2$  und terminiert nach endlich vielen Schritten wegen Lemma 3.8. Wegen (3.6) und  $\nabla f(x)^T d < 0$  gilt

$$\Psi(t) = f(x + td) - f(x) - \alpha t \nabla f(x)^T d \xrightarrow{t \rightarrow \infty} \infty.$$

Damit ist (3.5a) für  $t$  hinreichend groß nicht erfüllt. Schritt 4 terminiert daher auch nach endlich vielen Schritten. Zu Beginn von Schritt 5 gelten:

$$\underline{t} < \bar{t}, \quad t = \underline{t} \text{ erfüllt (3.5a)}, \quad t = \bar{t} \text{ erfüllt (3.5a) nicht} \quad (3.7)$$

Nun wird in jeder Iteration von Schritt 5 die Länge des Intervalls  $[\underline{t}, \bar{t}]$  halbiert, wobei (3.7) stets gilt. Hierbei wird entweder  $\underline{t}$  vergrößert oder  $\bar{t}$  verkleinert. Angenommen, Schritt 5 terminiert nicht nach endlich vielen Iterationen. Dann gibt es  $t^*$  mit

$$\underline{t} \rightarrow (t^*)^- \quad \bar{t} \rightarrow (t^*)^+$$

Wegen (3.6) gilt aus Stetigkeitsgründen  $\Psi(t^*) = 0$  (denn  $\Psi(\underline{t}) \leq 0$  und  $\Psi(\bar{t}) > 0$ ). Aus  $\Psi(\bar{t}) > 0 = \Psi(t^*)$  und  $\bar{t} \rightarrow (t^*)^+$  folgt nun  $\Psi'(\bar{t}) = \lim_{\bar{t} \rightarrow (t^*)^+} \underbrace{\frac{1}{\bar{t} - (t^*)^+}}_{>0} (\underbrace{\Psi(\bar{t})}_{>0} - \underbrace{\Psi(t^*)}_{=0}) \geq 0$ , also

$$\nabla f(x + t^*d)^T d \geq \alpha \nabla f(x)^T d > \varrho \nabla f(x)^T d.$$

Für  $\underline{t}$  hinreichend nahe bei  $t^*$  gilt dann aus Stetigkeitsgründen

$$\nabla f(x + \underline{t}d)^T d \geq \varrho \nabla f(x)^T d,$$

und (3.5b) ist daher an  $t = \underline{t}$  erfüllt, so dass die Iterationen in Schritt 5 abbrechen. Dies ist ein Widerspruch.  $\square$

Ohne Beweis geben wir den folgenden Satz an:

**Satz 3.17. (Existenz zulässiger Wolfe-Powell-Schrittweiten)**

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar und  $\alpha \in (0, \frac{1}{2})$ ,  $\rho \in [\alpha, 1)$  und  $x^0 \in \mathbb{R}^n$  fest vorgegeben. Zu  $x \in \mathcal{L}(x^0) = \{\tilde{x} \in \mathbb{R}^n \mid f(\tilde{x}) \leq f(x^0)\}$  und einer Richtung  $d \in \mathbb{R}^n$  mit  $\nabla f(x)^T d < 0$  sei

$$\mathcal{T}_{WP}(x, d) = \{t > 0 \mid (3.5a) \text{ und } (3.5b) \text{ sind erfüllt}\}.$$

Dann gelten:

1. Ist  $f$  nach unten beschränkt, so ist  $\mathcal{T}_{WP}(x, d) \neq \emptyset$ , d.h. die Wolfe-Powell-Schrittweitenstrategie ist wohldefiniert.
2. Ist weiter  $\nabla f$  auf  $\mathcal{L}(x^0)$  Lipschitz-stetig, so gibt es eine Konstante  $\theta > 0$  (unabhängig von  $x$  und  $d$ ) mit

$$f(x + td) \leq f(x) - \theta \left( \frac{\nabla f(x)^T d}{\|d\|} \right)^2 \quad \text{für alle } t \in \mathcal{T}_{WP}(x, d).$$

**Bemerkung 3.18.**

Die WOLFE-POWELL-Schrittweitenstrategie garantiert damit den in Satz 3.6 geforderten Abstieg.  $\diamond$

**3.2.3. Die strenge Wolfe-Powell-Regel**

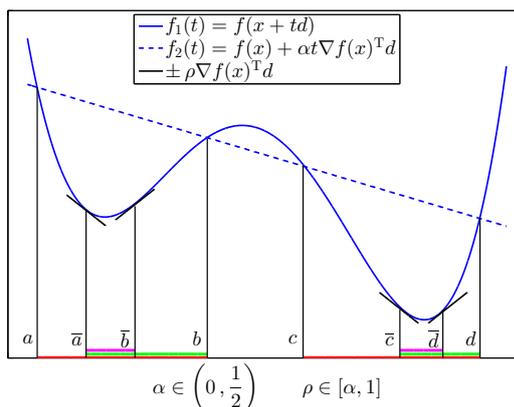
Seien  $\alpha \in (0, \frac{1}{2})$  und  $\rho \in [\alpha, 1)$  fest vorgegeben. Die **strenge Wolfe-Powell-Regel** lautet:  
Zu  $x, d \in \mathbb{R}^n$  mit  $\nabla f(x)^T d < 0$  bestimme eine Schrittweite  $t > 0$  mit

$$f(x + td) \leq f(x) + \alpha t \nabla f(x)^T d, \quad (3.8a)$$

$$|\nabla f(x + td)^T d| \leq -\rho \nabla f(x)^T d. \quad (3.8b)$$

Im Gegensatz zur klassischen Wolfe-Powell-Regel wird also nicht nur gefordert, dass der Graph von  $\varphi$  nicht zu steil abfällt, sondern auch, dass er nicht zu steil ansteigt.

Die strengen-textscWolfe-Powell-Bedingungen gelten für  $t \in [\bar{a}, \bar{b}]$  und für  $t \in [\bar{c}, \bar{d}]$ . Für sehr kleines  $\rho$  (und damit auch kleines  $\alpha$ ) ist eine Schrittweite, welche Bedingung (3.8b) erfüllt, nahe an einem stationären Punkt von  $\varphi$ .



**Fig. 6:** Die strenge Wolfe-Powell-Regel.

**3.3. Praktische Aspekte**

Die vorgestellten Schrittweitalgorithmen sind idealisiert. In der Praxis sind  $f$  und  $\nabla f$  nur maschinen- und/oder problemabhängig genau. Werden diese Ungenauigkeiten nicht berücksichtigt, führt dies schnell zu Endlosschleifen.

Ideal wäre es, wenn mit den Funktionswerten  $\varphi(t)$ ,  $\varphi(0)$  und den Ableitungen  $\varphi'(t)$ ,  $\varphi'(0)$  Fehlerschranken  $\varepsilon(t)$ ,  $\varepsilon(0)$  und  $\hat{\varepsilon}(t)$ ,  $\hat{\varepsilon}(0)$  mitgeliefert würden. Dann werden die beiden Bedingungen

$$\varphi(t) \leq \varphi(0) + \alpha t \varphi'(0) \quad \text{bzw.} \quad \varphi'(t) \geq \rho \varphi'(0)$$

ersetzt durch

$$\varphi(t) \leq \varphi(0) + \alpha t (\varphi'(0) + \hat{\varepsilon}(0)) + \varepsilon(0) + \varepsilon(t) \quad \text{bzw.} \quad \varphi'(t) \geq \rho (\varphi'(0) - \hat{\varepsilon}(0)) - \hat{\varepsilon}(t).$$

Weiter ist abzubrechen, wenn das Intervall  $[\underline{t}, \bar{t}]$  „zu klein“ wird, d.h. wenn  $\bar{t} - \underline{t}$  klein wird. Ferner sollte man stets eine untere Schranke für  $f$  mitführen.

### 3.4. Gradientenverfahren

Eine nahe liegende Wahl für  $d$  – auch in Hinblick auf die Winkelbedingung aus Satz 3.6 – ergibt sich als Lösung von

$$\min \nabla f(x)^T d \quad \text{u.d.N.} \quad \|d\| = 1.$$

Das Ziel ist also,  $d$  als jene Richtung zu bestimmen, in welche  $f$  in  $x$  am steilsten fällt. Gemäß der CAUCHY-SCHWARZschen Ungleichung wird dieses Problem gelöst von  $d = -\frac{\nabla f(x)}{\|\nabla f(x)\|}$ .

---

#### Algorithmus 3.19 (Gradientenverfahren, Verfahren des steilsten Abstiegs)

---

**Eingabe:** Startwert  $x^0$ ,  $\alpha \in (0, 1)$ ,  $\beta \in (0, 1)$ .

```

1: for  $k = 0, 1, \dots$  do
2:   if  $\nabla f(x^k) = 0$  then
3:     break
4:   else
5:     Setze  $d^k = -\nabla f(x^k)/\|\nabla f(x^k)\|$ .
6:     Bestimme die Schrittweite  $t_k$  gemäß Algorithmus 3.9 mit  $\nu_0 = \nu^0 = \beta$ .
7:     Setze  $x^{k+1} = x^k + t_k d^k$ .
8:   end if
9: end for

```

---

#### Bemerkung 3.20.

In der Implementation von Algorithmus 3.9 wird die Abbruchbedingung ersetzt durch

$$\|\nabla f(x^k)\| \leq \tau_{\text{rel}} \|\nabla f(x^0)\| + \tau_{\text{abs}}$$

mit für das vorliegende Problem gewählten relativen und absoluten Toleranz  $0 < \tau_{\text{rel}} \leq \tau_{\text{abs}}$ . Ferner wird die **for**-Schleife nach einer maximalen Anzahl von Iterationen  $k_{\text{max}}$  abgebrochen.  $\diamond$

Verwenden wir die WOLFE-POWELL-Schrittweitenstrategie, dann folgt sofort aus den Sätzen 3.6 und 3.17, dass jeder Häufungspunkt der Folge  $(x^k)_{k \in \mathbb{N}}$  ein stationärer Punkt von  $f$  ist. Eine analoge Aussage gilt auch für die strenge WOLFE-POWELL-Regel. Die Armijo-Bedingung hingegen erfüllt die Bedingung des hinreichenden Abstiegs nicht notwendigerweise.

Das folgende Resultat folgt auch aus Satz 3.12. Wir geben hier einen alternativen Beweis an.

#### Satz 3.21.

Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar. Dann terminiert der Algorithmus 3.19 entweder endlich mit einem stationären Punkt, oder er erzeugt eine endliche Folge  $(x^k)_{k \in \mathbb{N}}$  mit diesen Eigenschaften

1. Für alle  $k$  gilt  $f(x^{k+1}) < f(x^k)$ .
2. Jeder Häufungspunkt von  $(x^k)_{k \in \mathbb{N}}$  ist ein stationärer Punkt von  $f$

**Beweis.** (vergleiche [8, Satz 7.7])

Wir müssen nur den Fall betrachten, dass der Algorithmus nicht endlich abbricht. Nach Lemma 3.8 erzeugt, das Verfahren dann unendliche Folgen  $(x^k)_{k \in \mathbb{N}}$  und  $(t_k)_{k \in \mathbb{N}} \subset (0, 1]$  mit  $\nabla f(x^k) \neq 0$  und

$$f(x^{k+1}) - f(x^k) = f(x^k + t_k d^k) - f(x^k) \leq -\alpha t_k \|\nabla f(x^k)\|^2 < 0$$

und damit ist der 1. Teil gezeigt.

Nun zu Teil 2. Sei  $x^*$  ein Häufungspunkt von  $(x^k)_{k \in \mathbb{N}}$  und  $(x^{k_\nu})_{\nu \in \mathbb{N}}$  eine Teilfolge mit Grenzwert  $\lim_{\nu \rightarrow \infty} x^{k_\nu} = x^*$ . Die Folge  $(f(x^k))_{k \in \mathbb{N}}$  ist monoton fallend und besitzt daher einen Grenzwert  $z \in \mathbb{R} \cup \{-\infty\}$ . Daraus folgt insbesondere:  $(f(x^{k_\nu}))_{\nu \in \mathbb{N}}$  konvergiert gegen  $z$ . Wegen der Stetigkeit von  $f$  und wegen  $\lim_{\nu \rightarrow \infty} x^{k_\nu} = x^*$  folgt aber auch  $\lim_{\nu \rightarrow \infty} f(x^{k_\nu}) = f(x^*)$ . Aus der Eindeutigkeit des Grenzwertes bekommen

wir  $z = f(x^*)$  und  $\lim_{k \rightarrow \infty} f(x^k) = f(x^*)$ .

Aus der Armijo-Regel schließen wir

$$\infty > f(x^0) - f(x^*) = \sum_{k=0}^{\infty} (f(x^k) - f(x^{k+1})) \geq \alpha \sum_{k=0}^{\infty} t_k \|\nabla f(x^k)\|^2$$

Das bedeutet aber

$$t_k \|\nabla f(x^{k_\nu})\|^2 \rightarrow 0 \quad k \rightarrow \infty \quad (3.9)$$

Den Rest des Beweises führen wir per Widerspruch: Angenommen, es gilt  $\nabla f(x^*) \neq 0$ . Aus der Stetigkeit von  $\nabla f$  und aus  $\lim_{\nu \rightarrow \infty} x^{k_\nu} = x^*$  erhalten wir die Existenz eines  $N_1 \in \mathbb{N}$  mit

$$\|\nabla f(x^{k_\nu})\| \geq \frac{1}{2} \|\nabla f(x^*)\| > 0 \quad \text{für alle } \nu \geq N_1$$

Mit (3.9) gilt dann  $t_{k_\nu}$  für  $\nu \rightarrow \infty$ . Insbesondere gibt es ein  $N_2 \in \mathbb{N}$  mit  $N_2 \geq N_1$ , so dass  $t_{k_\nu} \leq \beta$  für  $\nu \geq N_2$  erfüllt ist. Aus der Armijo-Schrittweitenregel ergibt sich dann

$$f(x^{k_\nu} + \underbrace{\beta^{-1} t_{k_\nu}}_{> t_{k_\nu}} d^{k_\nu}) - f(x^{k_\nu}) > -\alpha \beta^{-1} t_{k_\nu} \|\nabla f(x^{k_\nu})\|^2 \quad \text{für alle } \nu \geq N_2 \quad (3.10)$$

Wir setzen  $s_{k_\nu} = \beta^{-1} t_{k_\nu}$ . Dann ist  $(s_{k_\nu})_{\nu \in \mathbb{N}}$  eine Nullfolge, da  $t_{k_\nu} \xrightarrow{\nu \rightarrow \infty} 0$ . Mit dem Mittelwertsatz folgt

$$\lim_{\nu \rightarrow \infty} \frac{1}{s_{k_\nu}} (f(x^{k_\nu} + s_{k_\nu} d^{k_\nu}) - f(x^{k_\nu})) = \lim_{\nu \rightarrow \infty} \frac{1}{s_{k_\nu}} (s_{k_\nu} \nabla f(x^{k_\nu} + t_{k_\nu} d^{k_\nu})^\top d^{k_\nu}) = -\|\nabla f(x^*)\|_2^2$$

mit Zwischenstellen  $t_{k_\nu} \in [0, s_{k_\nu}]$ . Ferner gilt

$$\lim_{\nu \rightarrow \infty} \|\nabla f(x^{k_\nu})\|_2^2 = \|\nabla f(x^*)\|_2^2.$$

Damit ergibt sich aus (3.10) der Widerspruch

$$-\|\nabla f(x^*)\|_2^2 = \lim_{\nu \rightarrow \infty} \frac{f(x^{k_\nu} + s_{k_\nu} d^{k_\nu}) - f(x^{k_\nu})}{\beta^{-1} t_{k_\nu}} \geq \lim_{\nu \rightarrow \infty} -\alpha \|\nabla f(x^{k_\nu})\|_2^2 = -\alpha \|\nabla f(x^*)\|_2^2$$

d.h.,  $0 \geq \underbrace{(1 - \alpha)}_{>0} \underbrace{\|\nabla f(x^*)\|_2^2}_{>0} > 0$ , was nicht möglich ist. Also ist die Annahme  $\nabla f(x^*) \neq 0$  falsch.  $\square$

### Bemerkung 3.22.

Das Konvergenzverhalten des steilsten Abstiegs kann sehr schlecht sein. Für eine *quadratische Form*  $f(x) = \frac{1}{2} x^\top A x - b^\top x + c$  mit positiv definiten Matrix  $A \in \mathcal{S}_n$ ,  $b \in \mathbb{R}^n$  und  $c \in \mathbb{R}$  lässt sich zeigen, dass

$$\|x^k - x^*\|_2 \leq \sqrt{\kappa_2(A)} \left( \frac{\kappa_2(A) - 1}{\kappa_2(A) + 1} \right)^k \|x^0 - x^*\|_2, \quad (3.11)$$

erfüllt ist, wobei  $\kappa_2(A) = \lambda_g(A)/\lambda_s(A)$  die spektrale Konditionszahl von  $A$  bezeichnet mit

$$\lambda_g = \max \{ |\lambda| \mid \lambda \text{ ist Eigenwert von } A \} \quad \text{und} \quad \lambda_s = \min \{ |\lambda| \mid \lambda \text{ ist Eigenwert von } A \}.$$

Eine mögliche Abhilfe für die langsame Konvergenz des Verfahrens des steilsten Abstiegs besteht darin,  $d^k := -H^{-1} \nabla f(x^k)$  mit geeigneter positiv definiten Matrix  $H \in \mathcal{S}_n$  zu setzen.  $H$  soll überdies so gewählt sein, dass gilt

$$0 < \frac{\lambda_g(H^{-1}A)}{\lambda_s(H^{-1}A)} = \kappa_2(H^{-1}A) < \frac{\lambda_g(A)}{\lambda_s(A)} = \kappa_2(A). \quad \diamond$$

## 4. Verfahren der konjugierten Gradienten

### 4.1. A-konjugierte Richtungen

Wir betrachten die Funktion

$$f(x) = \frac{1}{2}x^T Ax - b^T x + c$$

mit  $A \in \mathcal{S}_n$  positiv definit,  $b \in \mathbb{R}^n$  und  $c \in \mathbb{R}$ .

Das Verfahren des steilsten Abstiegs kann langsam sein, obwohl wir die exakte Minimierungsregel in der Schrittweitenwahl verwenden. Die Matrix  $A$  ist häufig groß und häufig dünn besetzt, bzw. speziell besetzt. Dann ist es oft einfacher, das Matrix-Vektor-Produkt  $Ax$  auszugeben. Das **CG-Verfahren** (CG – Conjugate Gradient) bricht nach endlich vielen Schritten ab und  $A$  wird nicht als Feld abgespeichert sondern nur  $Ax$  wird benötigt.

Die notwendige und hinreichende Bedingung für eine Minimalstelle  $x^*$  von

$$\min_{x \in \mathbb{R}^n} \left( \frac{1}{2}x^T Ax - b^T x + c \right)$$

lautet

$$Ax^* = b.$$

Gegeben sei ein Startpunkt  $x^0 \in \mathbb{R}^n$ . Wir wollen nun ein Verfahren angeben, das iterativ eine Lösung von  $Ax = b$  konstruiert. Dazu betrachten wir folgende motivierende Aussage:

#### Lemma 4.1.

Sei  $f(x) = \frac{1}{2}x^T Ax - b^T x + c$  mit  $A \in \mathcal{S}_n$  positiv definit,  $b \in \mathbb{R}^n$ ,  $c \in \mathbb{R}$  und  $x^0 \in \mathbb{R}^n$ . Seien weiter  $d^0, d^1, \dots, d^{n-1} \in \mathbb{R}^n$  mit

$$(d^i)^T Ad^j = 0 \quad \forall 1 \leq i \neq j \leq n-1. \quad (4.1)$$

Dann liefert das Verfahren der sukzessiven eindimensionalen Minimierung längs  $d^0, d^1, \dots, d^{n-1} \in \mathbb{R}^n$ , d.h. die Berechnung einer Folge  $(x^k)_{k \in \mathbb{N}}$  aus  $x^{k+1} := x^k + t_k d^k$  mit

$$f(x^k + t_k d^k) = \min_{t \in \mathbb{R}} f(x^k + t d^k) \quad k = 0, 1, \dots, n-1 \quad (4.2)$$

nach höchstens  $n$  Schritten mit  $x^n$  die Minimalstelle  $x^*$  von  $f$ . Weiter gelten für  $k = 0, 1, \dots, n-1$  mit  $g^k := Ax^k - b = \nabla f(x^k)$  die Gleichungen

$$t_k = \frac{-(g^k)^T d^k}{(d^k)^T Ad^k} \quad (4.3)$$

und

$$(g^{k+1})^T d^j = 0, \quad j = 0, 1, \dots, k. \quad (4.4)$$

#### Beweis.

Es gilt:

$$\begin{aligned} f(x^k + t d^k) &= \frac{1}{2}t^2 (d^k)^T Ad^k + t((x^k)^T Ad^k - b^T d^k) + c + \frac{1}{2}(x^k)^T Ax^k - b^T x^k \\ &= \frac{t^2}{2} (d^k)^T Ad^k + t(g^k)^T d^k + c + \underbrace{\frac{1}{2}(x^k)^T Ax^k - b^T x^k}_{\text{unabhängig von } t} \end{aligned}$$

Aus (4.2) folgt, dass die Minimalstelle  $t_k$

$$t_k (d^k)^T Ad^k + (g^k)^T d^k = 0$$

erfüllt. Dies beweist (4.3). Weiter folgt, dass

$$\begin{aligned} 0 &= t_k(d^k)^T Ad^k + (g^k)^T d^k = (t_k(d^k)^T A + (g^k)^T A - b^T) d^k \\ &= (A(x^k + t_k d^k) - b)^T d^k = (Ax^{k+1} - b)^T d^k \\ &= (g^{k+1})^T d^k \quad \text{für } k = 0, 1, \dots, n-1 \end{aligned} \quad (4.5)$$

Verwendung von (4.1) für  $i \neq j$  ergibt

$$(g^{i+1} - g^i)^T d^j = (Ax^{i+1} - b - Ax^i + b)^T d^j = t_i(d^i)^T Ad^j = 0$$

Das liefert zusammen mit (4.5) für  $j = 0, 1, \dots, k \leq n-1$

$$(g^{k+1})^T d^j = \underbrace{(g^{j+1})^T d^j}_{=0 \text{ (4.5)}} + \underbrace{\sum_{i=j+1}^k (g^{i+1} - g^i)^T d^j}_{=0 \text{ da } i \neq j} = 0$$

woraus (4.4) folgt. Die Vektoren  $d^0, \dots, d^{n-1}$  sind wegen (4.1) bezüglich des Skalarproduktes

$$\langle u, v \rangle_A = u^T Av$$

paarweise orthogonal und linear unabhängig. Aus (4.4) folgt dann sofort, dass  $g^n = 0$  oder äquivalent  $x^n = x^*$ .  $\square$

#### Bemerkung 4.2.

Die Vektoren  $d^0, d^1, \dots, d^{n-1}$  mit der Eigenschaft (4.1) heißen  **$A$ -orthogonal** oder  **$A$ -konjugiert**.  $\diamond$

## 4.2. Bestimmung $A$ -konjugierter Richtungen

Wir starten mit

$$d^0 = -\nabla f(x^0) = -g^0$$

Angenommen, es sind bereits  $l+1$  Vektoren  $d^0, d^1, \dots, d^l$ , die  $A$ -konjugiert sind, d.h.

$$(d^i)^T Ad^j = 0 \quad \text{für } i, j = 0, 1, \dots, l \quad i \neq j \quad (4.6)$$

Nach Lemma 4.1 gelten (4.3) und (4.4) für  $k = 0, \dots, l$ . Wir nehmen an, dass  $g^{l+1} \neq 0$ . Wir verwenden den Ansatz

$$d^{l+1} := -g^{l+1} + \sum_{i=0}^l \beta_i^l d^i \quad (4.7)$$

(wegen (4.4) ist  $g^{l+1}$  linear unabhängig von  $d^i$  für  $i = 0, \dots, l$ ) Es soll

$$(d^{l+1})^T Ad^j = 0 \quad \text{für } j = 0, \dots, l$$

gelten. Wir haben:

$$\begin{aligned} (d^{l+1})^T Ad^j &= \left( -g^{l+1} + \sum_{i=0}^l \beta_i^l d^i \right)^T Ad^j \\ &= -(g^{l+1})^T Ad^j + \sum_{i=0}^l \beta_i^l (d^i)^T Ad^j = -(g^{l+1})^T Ad^j + \beta_j^l (d^j)^T Ad^j, \end{aligned}$$

da nach (4.6)  $(d^i)^T Ad^j = 0$  für  $i \neq j$ . Daraus folgt

$$\beta_j^l = \frac{(g^{l+1})^T Ad^j}{(d^j)^T Ad^j} \quad \text{für } j = 0, \dots, l \quad (4.8)$$

Wir wollen noch weitere Eigenschaften herleiten: Multipliziert man (4.7) von links mit  $(g^{l+1})^T$  so gilt

$$(g^{l+1})^T \left( -g^{l+1} + \sum_{i=0}^l \beta_i d^i \right) = -\|g^{l+1}\|^2 + \sum_{i=0}^l \beta_i (g^{l+1})^T d^i = -\|g^{l+1}\|^2 < 0.$$

Offensichtlich ist  $d^{l+1}$  eine Abstiegsrichtung für  $f$  in  $x^{l+1}$ . Aus (4.3) folgt

$$t_{l+1} = \frac{-(g^{l+1})^T d^{l+1}}{(d^{l+1})^T A d^{l+1}} > 0.$$

Aufgrund der Konstruktion der  $d^k$ ,  $k = 0, \dots, l$ , gilt weiter, dass

$$(g^k)^T d^k = -\|g^k\|^2 < 0 \quad \text{und} \quad t_k > 0 \quad \forall k = 0, \dots, l.$$

Eine weitere Orthogonalitätsbedingung ergibt sich folgendermaßen:

$$\begin{aligned} (g^{l+1})^T g^j &= (g^{l+1})^T \left( \sum_{i=0}^{j-1} \beta_i^{j-1} d^i - d^j \right) \\ &= \sum_{i=0}^{j-1} \beta_i^{j-1} \underbrace{(g^{l+1})^T d^i}_{=0} - \underbrace{(g^{l+1})^T d^j}_{=0} = 0 \quad \text{für} \quad j = 0, \dots, l. \end{aligned} \quad (4.9)$$

Also gilt für die rechte Seite von (4.8) wegen

$$g^{j+1} - g^j = Ax^{j+1} - Ax^j = t_j A d^j \quad \forall j = 0, \dots, l$$

die Gleichheit

$$(g^{l+1})^T A d^j = \frac{1}{t_j} (g^{l+1})^T (g^{j+1} - g^j) = 0 \quad \forall j = 0, \dots, l-1.$$

Daraus folgt:

$$\beta_j^l = \frac{(g^{l+1})^T A d^j}{(d^j)^T A d^j} = 0 \quad \text{für} \quad j = 0, \dots, l-1$$

und

$$\beta_l^l = \frac{(g^{l+1})^T A d^l}{(d^l)^T A d^l} = \frac{1}{t_l} \frac{(g^{l+1})^T (g^{l+1} - g^l)}{(d^l)^T A d^l} = \frac{\|g^{l+1}\|^2}{(d^l)^T (g^{l+1} - g^l)} = \frac{\|g^{l+1}\|^2}{-(d^l)^T g^l} = \frac{\|g^{l+1}\|^2}{\|g^l\|^2} =: \beta_l$$

Somit reduziert sich (4.7) auf

$$d^{l+1} = -g^{l+1} + \beta_l d^l.$$

Weiter sei bemerkt, dass wegen  $g^k = Ax^k - b$  gilt:

$$g^{k+1} - g^k = Ax^{k+1} - Ax^k = t_k A d^k$$

Also kann  $g^k$  von Schritt zu Schritt aufaddiert werden, ohne dass eine weitere Matrix-Vektor-Multiplikation notwendig ist, da  $A d^k$  bereits in der Bestimmung von  $t_k$  aufgetaucht ist. Eine Skalarproduktauswertung erspart man sich, wenn man (4.3) mit Hilfe von  $(g^k)^T d^k = -\|g^k\|^2$  umformt zu

$$t_k = \frac{\|g^k\|^2}{(d^k)^T A d^k}.$$

Der **CG-Algorithmus** lautet nun:

#### Bemerkung 4.4.

1. Der Hauptaufwand liegt in den Berechnungen von  $A d^k$ . Da dieses Produkt zweimal benötigt wird, sollte man selbiges als  $z^k := A d^k$  abspeichern.

**Algorithmus 4.3 (CG-Algorithmus für quadratische Minimierung)****Eingabe:**  $x^0 \in \mathbb{R}^n$ 

- 1: Setze  $g^0 = Ax^0 - b$ ,  $d^0 = -g^0$ ,  $k = 0$  und wähle  $\epsilon > 0$
- 2: **while**  $\|g^k\| \geq \epsilon$  **do**
- 3: Bestimme Schrittweite  $t_k = \frac{\|g^k\|^2}{(d^k)^T A d^k}$
- 4: Bestimme nächste Iterierte  $x^{k+1} = x^k + t_k d^k$
- 5: Setze  $g^{k+1} = g^k + t_k A d^k$
- 6: Setze  $\beta_k = \frac{\|g^{k+1}\|^2}{\|g^k\|^2}$
- 7: Setze  $d^{k+1} = -g^{k+1} + \beta_k d^k$
- 8:  $k = k + 1$
- 9: **end while**

2. Wegen (4.9) kann  $\beta_k$  gemäß

$$\beta_k = \frac{(g^{k+1} - g^k)^T g^{k+1}}{\|g^k\|^2} \quad (4.10)$$

berechnet werden. Der Vorteil von (4.10): Die Richtungen sind in der numerischen Praxis schnell nicht mehr  $A$ -konjugiert. Dann kann zum Beispiel  $t_k$  sehr klein oder überhaupt negativ werden. Es empfiehlt sich zum Beispiel  $t_k := \max\{0, \frac{\|g^k\|^2}{(d^k)^T A d^k}\}$  zu wählen. Angenommen  $t_k \approx 0$ : Es ergibt sich aus  $g^{k+1} = g^k + t_k A d^k$  mit (4.10) also  $\beta_k \approx 0$  und daher

$$d^{k+1} \approx -g^{k+1}.$$

Es wird daher unter Verwendung von (4.10) ein automatischer **Restart** mit der Richtung des steilsten Abstiegs durchgeführt.

3. Besitzt  $A \in \mathbb{R}^{n \times n}$  genau  $m \leq n$  paarweise verschiedene Eigenwerte, so bricht das CG-Verfahren nach  $m$  Schritten ab. Ferner bricht das Verfahren nach  $m$  Schritten ab, wenn  $b$  als Linearkombination von höchstens  $m$  Eigenvektoren darstellbar ist und mit  $x^0 = 0 \in \mathbb{R}^n$  initialisiert wird.

◇

Sei  $\kappa_2(Q) = \lambda_{\max}(A)/\lambda_{\min}(A)$ , dann gilt für die Folge  $(x^k)_{k \in \mathbb{N}}$ :

$$\|x^k - x^*\| \leq 2\sqrt{\kappa_2(A)} \left( \frac{\sqrt{\kappa_2(A)} - 1}{\sqrt{\kappa_2(A)} + 1} \right)^k \|x^0 - x^*\|;$$

vergleiche Abschätzung (3.11) für das Gradientenverfahren. Die Absicht besteht nun darin, die Matrix  $A$  so zu transferieren, dass möglichst viele Eigenwerte der transformierten Matrix 1 sind (und die restlichen nahe bei 1). Sei  $W \in \mathbb{R}^{n \times n}$  positiv definit und symmetrisch. Die Lösung von  $Ax = b$  kann gefunden werden, indem man das System

$$W^{-1/2} A W^{-1/2} y = W^{-1/2} b$$

löst und  $x = W^{-1/2} y$  verwendet. Die Matrix  $R = W^{-1/2} A W^{-1/2}$  besitzt dieselben Eigenwerte wie  $W^{-1} A$ , da  $W^{-1/2} R W^{1/2} = W^{-1} A$ . Der Wunsch, die Matrix  $W$  so zu bestimmen, dass möglichst viele Eigenwerte eins sind, entspricht der Tatsache, dass  $\kappa_2(W^{-1} A)$  möglichst klein ist. Die Matrix  $W$  nennt man **Präkonditionierungsmatrix** und nennt das Verfahren **Präkonditionierung**. In der Praxis verwendet man  $W$  und nicht  $W^{-1/2}$ . In der numerischen Praxis wird natürlich nicht  $W^{-1}$  gebildet, sondern es wird  $Wd = g$  gelöst. Die Umsetzung des vorkonditionierten CG-Verfahrens ist in Algorithmus 4.5 beschrieben.

**5. Newton-Verfahren****5.1. Konvergenzraten und das lokale Newton-Verfahren**

**Algorithmus 4.5 (Präkonditionierter CG-Algorithmus)****Eingabe:**  $W \in \mathcal{S}_n$  positiv definit  $x^0 \in \mathbb{R}^n$ 

- 1: Setze  $g^0 = Ax^0 - b$ ,  $d^0 = -W^{-1}g^0$ ,  $k = 0$  und wähle  $\epsilon > 0$
- 2: **while**  $\|g^k\| \geq \epsilon$  **do**
- 3: Bestimme Schrittweite  $t_k = \frac{(g^k)^T W^{-1} g^k}{(d^k)^T A d^k}$
- 4: Bestimme nächste Iterierte  $x^{k+1} = x^k + t_k d^k$
- 5: Setze  $g^{k+1} = g^k + t_k A d^k$
- 6: Setze  $\beta_k = \frac{(g^{k+1})^T W^{-1} g^{k+1}}{(g^k)^T W^{-1} g^k}$
- 7: Setze  $d^{k+1} = -W^{-1} g^{k+1} + \beta_k d^k$
- 8:  $k = k + 1$
- 9: **end while**

**Definition 5.1.**Seien  $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  und  $x, x^* \in \mathbb{R}^n$  geben.

1.  $f(x) = o(g(x))$  für  $x \rightarrow x^*$ , falls eine (uniforme) Konstante  $C > 0$  und eine Umgebung  $U$  von  $\bar{x}$  existieren, so dass gilt

$$\|f(x)\| \leq C \|g(x)\| \quad \text{für alle } x \in U \setminus \{x^*\}.$$

2.  $f(x) = \mathcal{O}(g(x))$  für  $x \rightarrow x^*$ , falls für alle  $\epsilon > 0$  eine Umgebung  $U_\epsilon$  von  $x^*$  existiert mit

$$\|f(x)\| \leq \epsilon \|g(x)\| \quad \text{für alle } x \in U_\epsilon \setminus \{x^*\}.$$

**Definition 5.2. (Konvergenzraten)**Die Folge  $(x^k) \subset \mathbb{R}^n$ 

1. konvergiert **q-linear** mit Rate  $0 < \gamma < 1$  gegen  $\bar{x} \in \mathbb{R}^n$ , falls es ein  $l \geq 1$  gibt, so dass gilt:

$$\|x^{k+1} - x^*\| \leq \gamma \|x^k - x^*\| \quad \forall k \geq l.$$

2. konvergiert **q-superlinear** gegen  $x^* \in \mathbb{R}^n$ , falls  $x^k \rightarrow x^*$  gilt sowie

$$\|x^{k+1} - x^*\| \leq o(\|x^k - x^*\|) \quad \text{für } k \rightarrow \infty.$$

Die Bedingung ist gleichbedeutend mit

$$\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} \rightarrow 0 \quad \text{für } k \rightarrow \infty.$$

3. konvergiert **q-quadratisch** gegen  $x^* \in \mathbb{R}^n$ , falls  $x^k \rightarrow x^*$  gilt sowie,

$$\|x^{k+1} - x^*\| \leq \mathcal{O}(\|x^k - x^*\|^2) \quad \text{für } k \rightarrow \infty.$$

Die Bedingung ist gleichbedeutend damit, dass es  $C \geq 0$  gibt mit

$$\|x^{k+1} - x^*\| \leq C \|x^k - x^*\|^2 \quad \forall k \geq 0.$$

**Bemerkung 5.3.**In der Literatur gibt es ferner die Begriffe  $r$ -linear,  $r$ -superlinear und  $r$ -quadratisch

◇

**Satz 5.4.**

Sei  $(x^k)_{k \in \mathbb{N}}$  eine Folge in  $\mathbb{R}^n$  mit  $\lim x^k = x^* \in \mathbb{R}^n$ . Dann folgt

$$0 \leq \left| 1 - \frac{\|x^{k+1} - x^k\|}{\|x^k - x^*\|} \right| \leq \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} \quad \text{für } x^k \neq x^* \quad (5.1)$$

Wenn  $(x^k)_{k \in \mathbb{N}}$  superlinear gegen  $x^*$  konvergiert, mit  $x^k \neq x^*$  für alle  $k \geq k_0$ , dann gilt

$$\lim_{k \rightarrow \infty} \frac{\|x^{k+1} - x^k\|}{\|x^k - x^*\|} = 0 \quad (5.2)$$

**Beweis.**

Die erste Ungleichung ist trivial. Für die rechte Ungleichung nehmen wir zunächst ein positives (und gleich null) Argument im Betrag an. Es gilt

$$\begin{aligned} \|x^k - x^*\| &= \|-x^{k+1} + x^k + x^{k+1} - x^*\| \\ &\leq \|x^{k+1} - x^k\| + \|x^{k+1} - x^*\| \end{aligned} \quad (5.3)$$

Wegen (5.3) erhalten wir

$$\|x^k - x^*\| - \|x^{k+1} - x^k\| \leq \|x^{k+1} - x^*\|,$$

woraus durch Division mit  $\|x^k - x^*\| > 0$  bereits die Behauptung folgt

$$1 - \frac{\|x^{k+1} - x^k\|}{\|x^k - x^*\|} \leq \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|}.$$

Nehmen wir nun ein strikt negatives Argument im Betrag an. Es gilt

$$\begin{aligned} \|x^{k+1} - x^k\| &= \|x^{k+1} - x^* - x^k + x^*\| \\ &\leq \|x^{k+1} - x^*\| + \|x^k - x^*\| \end{aligned} \quad (5.4)$$

Wegen (5.4) erhalten wir

$$-\|x^k - x^*\| + \|x^{k+1} - x^k\| \leq \|x^{k+1} - x^*\|,$$

woraus durch Division mit  $\|x^k - x^*\| > 0$  bereits die Behauptung folgt

$$-1 + \frac{\|x^{k+1} - x^k\|}{\|x^k - x^*\|} \leq \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|}.$$

Bilden wir den Grenzwert dieser Ungleichungskette unter der Voraussetzung der Superlinearität so gilt:

$$0 = \lim_{k \rightarrow \infty} 0 \leq \lim_{k \rightarrow \infty} \left| 1 - \frac{\|x^{k+1} - x^k\|}{\|x^k - x^*\|} \right| \leq \lim_{k \rightarrow \infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 0$$

und damit  $\lim_{k \rightarrow \infty} \frac{\|x^{k+1} - x^k\|}{\|x^k - x^*\|} = 1$  □

Wir charakterisieren superlineare Konvergenz ohne Beweis; vergleiche dazu [8, Satz 11.2]

**Satz 5.5.**

Seien  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  stetig differenzierbar und  $x^* \in \mathbb{R}^n$  ein Punkt, in dem  $F'(x^*)$  regulär ist. Sei  $(x^k)_{k \in \mathbb{N}} \subset \mathbb{R}^n$  eine gegen  $x^*$  konvergente Folge mit  $x_k \neq x^*$  für alle  $k \geq k_0$ . Dann sind folgende Aussagen äquivalent:

1.  $(x^k)_{k \in \mathbb{N}}$  konvergiert superlinear gegen  $x^*$ , und es gilt  $F(x^*) = 0$
2.  $\|F(x^k) + F'(x^*)(x^{k+1} - x^k)\| = o(\|x^{k+1} - x^k\|)$
3.  $\|F(x^k) + F'(x^k)(x^{k+1} - x^k)\| = o(\|x^{k+1} - x^k\|)$

Sei die Folge  $(x^k)_{k \in \mathbb{N}}$  durch  $x^{k+1} = x^k + d^k$  gegeben mit  $d^k = -(H^k)^{-1}F(x^k)$ , wobei  $(H^k)_{k \in \mathbb{N}}$  eine Folge invertierbarer Matrizen in  $\mathbb{R}^{n \times n}$  ist. Dann folgt:

$$\|(H^k - F'(x^k))(x^{k+1} - x^k)\| = \|F(x^k) + F'(x^k)(x^{k+1} - x^k)\|,$$

was der 2. Aussage von Satz 5.5 entspricht. Wir erhalten auf diese Weise die DENNIS-MORÉ-Bedingung, für die Folge  $(x^k)_{k \in \mathbb{N}}$ : Gilt  $\lim_{k \rightarrow \infty} x^k = x^*$  mit  $\det(F'(x^*)) \neq 0$ , so sind folgende Aussagen äquivalent:

1.  $(x^k)_{k \in \mathbb{N}}$  konvergiert superlinear gegen  $x^*$ , und es gilt  $F(x^*) = 0$
2.  $\|(H^k - F'(x^*))(x^{k+1} - x^k)\| = o(\|x^{k+1} - x^k\|)$
3.  $\|(H^k - F'(x^k))(x^{k+1} - x^k)\| = o(\|x^{k+1} - x^k\|)$

---

### Algorithmus 5.6 (lokales Newton-Verfahren für Gleichungssysteme $F(x) = 0$ )

---

**Eingabe:** Startpunkt  $x^0 \in \mathbb{R}^n$

- 1: **for**  $k = 0, 1, 2, \dots$  **do**
  - 2:   **if**  $F(x^k) = 0$  **then**
  - 3:     **stop for** und erhalte die Lösung  $\bar{x} := x^k \in \mathbb{R}^n$
  - 4:   **else**
  - 5:     Berechne den Newtonschritt  $d^k \in \mathbb{R}^n$  durch Lösen der Newtongleichung  $F'(x^k)d^k = -F(x^k)$
  - 6:     Setze  $x^{k+1} = x^k + d^k$
  - 7:   **end if**
  - 8: **end for**
- 

## 5.2. Das lokale Newton-Verfahren in der Optimierung

Wir setzen ab nun voraus, dass  $f$  und die lokale Minimalstelle  $x^*$  von  $f$  folgende Voraussetzungen erfüllen:

$$f \text{ ist zweimal stetig differenzierbar.} \quad (5.5a)$$

$$\exists \gamma > 0 : \forall x, y \in \mathbb{R}^n : \|\nabla^2 f(x) - \nabla^2 f(y)\| \leq \gamma \|x - y\|. \quad (5.5b)$$

$$x^* \text{ ist ein stationärer Punkt von } f: \nabla f(x^*) = 0. \quad (5.5c)$$

$$\nabla^2 f(x^*) \text{ ist positiv definit.} \quad (5.5d)$$

Zur Vereinfachung der Schreibweise bezeichne  $x_a$  den aktuellen Iterationspunkt und  $x_+$  die neue Iterierte. Wir betrachten ein **quadratisches Modell** von  $f$  um  $x_a$ :

$$m_a(x) = f(x_a) + \nabla f(x_a)^T(x - x_a) + \frac{1}{2}(x - x_a)^T \nabla^2 f(x_a)(x - x_a)$$

$$\nabla m_a(x) = \nabla f(x_a) + \nabla^2 f(x_a)(x - x_a)$$

Wenn  $\nabla^2 f(x_a)$  positiv definit ist, dann definieren wir  $x_+$  als die (eindeutige) Minimalstelle dieses Modells. Es gilt  $0 = \nabla m_a(x_+) = \nabla f(x_a) + \nabla^2 f(x_a)(x_+ - x_a)$ . Umformen liefert die **Iterationsvorschrift des Newton-Verfahrens**, d.h.

$$x_+ = x_a - \nabla^2 f(x_a)^{-1} \nabla f(x_a). \quad (5.6)$$

Natürlich wird bei der numerischen Implementierung nicht  $\nabla^2 f(x_a)^{-1}$  berechnet, sondern es wird in jedem Iterationsschritt  $\nabla^2 f(x_a)d = -\nabla f(x_a)$  gelöst und  $x_+ = x_a + d$  gesetzt.

Falls  $x_a$  weit von einer lokalen Minimalstelle  $x^*$ , die 5.5 erfüllt, entfernt ist, dann kann  $\nabla^2 f(x_a)$  negative Eigenwerte haben. Also kann  $x_+$  eine lokale Minimalstelle oder ein Sattelpunkt sein. Um dies zu vermeiden, müssen geeignete Modifikationen eingeführt werden. Zunächst sei aber vorausgesetzt, dass  $x_a$  hinreichend nahe an  $x^*$  ist.

**Satz 5.7. (Quadratische Konvergenz des lokalen Newton-Verfahrens)**

1. Sei (5.5) erfüllt. Dann existieren Konstanten  $C > 0$  und  $\delta > 0$  derart, dass für alle  $x_a$  aus der Menge  $B(x^*, \delta) = \{x \in \mathbb{R}^n \mid \|x - x^*\| < \delta\}$  der **Newton-Schritt** (5.6) folgende Abschätzung erfüllt:

$$\|x_+ - x^*\| \leq C \|x_a - x^*\|^2.$$

2. Es sei (5.5) erfüllt. Dann existieren ein Radius  $\delta > 0$  und eine Konstante  $C > 0$ , so dass das Newton-Verfahren  $x^{k+1} = x^k - \nabla^2 f(x^k)^{-1} \nabla f(x^k)$  für  $x^0 \in B(x^*, \delta)$  gegen  $x^*$  konvergiert mit

$$\|x^{k+1} - x^*\| \leq C \|x^k - x^*\|^2.$$

**Beweis.**

1. Wähle  $\delta > 0$  derart, dass für alle  $x \in B(x^*, \delta)$  gilt

$$\|(\nabla^2 f(x))^{-1}\| \leq 2 \|(\nabla^2 f(x^*))^{-1}\|.$$

Da sich die Differenz zwischen  $x_+$  und  $x^*$  darstellen lässt als

$$\begin{aligned} x_+ - x^* &= x_a - x^* - (\nabla^2 f(x_a))^{-1} \nabla f(x_a) = (\nabla^2 f(x_a))^{-1} (\nabla^2 f(x_a)(x_a - x^*) - \nabla f(x_a)) \\ &= (\nabla^2 f(x_a))^{-1} \int_0^1 (\nabla^2 f(x_a) - \nabla^2 f(x^* + t(x_a - x^*))) (x_a - x^*) dt, \end{aligned}$$

erhalten wir daraus

$$\begin{aligned} \|x_+ - x^*\| &\leq \|(\nabla^2 f(x_a))^{-1}\| \int_0^1 \|(\nabla^2 f(x_a) - \nabla^2 f(x^* + t(x_a - x^*)))\| dt \cdot \|x_a - x^*\| \\ &\leq 2 \|(\nabla^2 f(x^*))^{-1}\| \gamma \int_0^1 \|x_a - x^* - t(x_a - x^*)\| dt \cdot \|x_a - x^*\| \\ &= 2\gamma \|(\nabla^2 f(x^*))^{-1}\| \cdot \|x_a - x^*\|^2 \int_0^1 1 - t dt = \underbrace{\gamma \|(\nabla^2 f(x^*))^{-1}\|}_{=C} \cdot \|x_a - x^*\|^2. \end{aligned}$$

2. Wähle  $\tilde{\delta} = \min(\delta, 1/C)$ . Dann gilt für jedes  $x^k \in B(x^*, \tilde{\delta})$ , dass auch  $x^{k+1}$  in  $B(x^*, \tilde{\delta})$  liegt, denn

$$\|x^{k+1} - x^*\| \leq C \|x^k - x^*\|^2 < C \tilde{\delta} \|x^k - x^*\| \leq \|x^k - x^*\| < \tilde{\delta}.$$

Aus  $x^0 \in B(x^*, \tilde{\delta})$  folgt also die Wohldefiniertheit des Newton-Verfahrens und  $(x^k)_{k \in \mathbb{N}} \subseteq B(x^*, \tilde{\delta})$ .  $\square$

**Bemerkung 5.8. (Terminierungsbedingung des lokalen Newton-Verfahrens)**

Ein natürliches Abbruchkriterium für das Newton-Verfahren (wie auch für die Gradientenverfahren des letzten Abschnitts) setzt sich aus einer **relativen** und einer **absoluten Fehlerschranke** zusammen: Sei  $\tau_{\text{rel}} \in (0, 1)$  eine erwünschte Reduktion in der Gradientennorm und  $\tau_{\text{abs}}$  mit  $1 \gg \tau_{\text{abs}} \geq \tau_{\text{rel}} > 0$  eine absolute Fehlerschranke, dann stoppt man das Verfahren, wenn

$$\|\nabla f(x^k)\| \leq \tau_{\text{rel}} \|\nabla f(x^0)\| + \tau_{\text{abs}}$$

erfüllt ist. Ist  $\|\nabla f(x^0)\|$  groß, so ist  $\tau_{\text{rel}} \|\nabla f(x^0)\|$  der dominante Term. Ist hingegen  $\|\nabla f(x^0)\|$  klein, dann dominiert  $\tau_{\text{abs}}$  die rechte Seite der obigen Ungleichung.  $\diamond$

**Bemerkung 5.9. (Approximationsfehler diskreter Ableitungen)**

Wir nehmen nun an, dass  $f$  nur approximativ ausgewertet werden kann. Betrachte die Einfachheit halber die Raumdimension  $n = 1$ , dann ist

$$\tilde{f}(x) = f(x) + \tilde{\varepsilon}_f(x) \quad \text{mit } \tilde{\varepsilon}_f(x) \geq 0 \text{ und } |\tilde{\varepsilon}_f(x)| \leq \varepsilon_f \text{ für ein } \varepsilon_f > 0.$$

Bestimmen wir nun die Ableitungen numerisch, z.B. durch Vorwärtsdifferenzen, so ergibt sich

$$D_h^+ f(x) = \frac{\tilde{f}(x+h) - \tilde{f}(x)}{h}.$$

Wir schätzen die Ordnung der Differenz zwischen  $D_h^+ f$  und  $f'$  ab: Nach dem Satz von Taylor existiert eine Zwischenstelle  $\xi \in (x, x+h)$  mit

$$\begin{aligned} \|D_h^+ f(x) - f'(x)\| &= \left\| \frac{\tilde{f}(x+h) - \tilde{f}(x)}{h} - f'(x) \right\| \\ &= \left\| \frac{f(x+h) + \tilde{\varepsilon}_f(x+h) - f(x) - \tilde{\varepsilon}_f(x)}{h} - f'(x) \right\| \\ &\leq \left\| \frac{f(x+h) - f(x)}{h} - f'(x) \right\| + \frac{2\varepsilon_f}{h} \\ &\leq \left\| \frac{f(x) + f'(x)h + \frac{1}{2}f''(\xi)h^2 - f(x)}{h} - f'(x) \right\| + \frac{2\varepsilon_f}{h} \\ &= \frac{h}{2} \|f''(\xi)\| + \frac{2\varepsilon_f}{h} = \mathcal{O}\left(h + \frac{\varepsilon_f}{h}\right). \end{aligned}$$

Die Minimalstelle  $h^*$  der Fehlerfunktion  $\text{err}_+(h) = h + \frac{\varepsilon_f}{h}$  erfüllt  $\text{err}'_+(h^*) = 1 - \frac{\varepsilon_f}{(h^*)^2} = 0$ . In der Tat ist  $h^* = \sqrt{\varepsilon_f}$  Minimalstelle der Fehlerfunktion. Der Fehler  $\varepsilon_g$  in der Ableitung ist also von der Ordnung  $\mathcal{O}(h^*) = \mathcal{O}(\sqrt{\varepsilon_f})$ .

Verwenden wir nun abermals Vorwärtsdifferenzen zur Approximation der Hessematrix, so ergibt sich für den Fehler  $\varepsilon_H$  die Größenordnung  $\mathcal{O}(\sqrt{\varepsilon_g}) = \mathcal{O}(\varepsilon_f^{1/4})$ . Dies impliziert, dass Hessematrizen, basierend auf zweifacher numerischer Differenzierung, relativ ungenau sind – selbst wenn  $\varepsilon_f = 10^{-16}$  die Größenordnung des **Maschinen-Epsilons** erreicht, folgt nur  $\varepsilon_H \approx 10^{-4}$ .

Im Falle **zentraler Differenzenapproximationen** ergibt sich ein besseres Ergebnis:  $\varepsilon_H = \mathcal{O}(\varepsilon_f^{4/9})$ . Betragen die Fehler in der ersten Ableitung  $\varepsilon_f = 10^{-16}$ , so erhalten wir so immerhin  $\varepsilon_H \approx 10^{-7.1}$ .

Konvergenz des Newton-Verfahrens ist nur zu erwarten, wenn  $\varepsilon_g \rightarrow 0$  im Laufe der Iteration gilt.  $\diamond$

**Satz 5.10. (Konvergenz des lokalen Newton-Verfahrens bei Diskretisierungsfehlern)**

Es seien die Bedingungen (5.5) erfüllt. Dann existieren Konstanten  $\bar{K} > 0$ ,  $\delta > 0$  und ein  $\varepsilon > 0$ , so dass für  $x_a \in B(x^*, \delta)$  und  $\|\varepsilon_H(x_a)\| < \varepsilon$  gilt

$$x_+ = x_a - (\nabla^2 f(x_a) + \varepsilon_H(x_a))^{-1} (\nabla f(x_a) + \varepsilon_g(x_a))$$

ist wohldefiniert, d.h.  $(\nabla^2 f(x_a) + \varepsilon_H(x_a))$  ist regulär, und erfüllt

$$\|x_+ - x^*\| \leq \bar{K} (\|x_a - x^*\|^2 + \underbrace{\|\varepsilon_H(x_a)\|}_{\text{beeinflusst Konv.geschw.}} \|x_a - x^*\| + \underbrace{\|\varepsilon_g(x_a)\|}_{\text{Genauigkeit}}).$$

**Bemerkung 5.11.**

Wir betrachten das **Newton-ähnliche Verfahren**

$$x^{k+1} = x^k - \nabla^2 f(x^0)^{-1} \nabla f(x^k), \quad x^0 \text{ Startwert, } k = 0, 1, \dots \quad (5.7)$$

Im Gegensatz zum klassischen Newton-Verfahren ist es hier nicht notwendig, in jedem Schritt die Hessematrix  $\nabla^2 f(x^k)$  aufzustellen. Es gelten

$$\varepsilon_H(x^k) = \nabla^2 f(x^0) - \nabla^2 f(x^k), \quad \|\varepsilon_H(x^k)\| \leq \varepsilon_H, \quad (5.8a)$$

$$\|\varepsilon_H(x^k)\| = \|\nabla^2 f(x^0) - \nabla^2 f(x^k)\| \leq \gamma \|x^0 - x^k\| \leq \gamma (\|x^0 - x^*\| + \|x^* - x^k\|). \quad (5.8b)$$

Die Konvergenz des Verfahrens (5.7) folgt aus Satz 5.10 mit  $\varepsilon_g = 0$  und  $\varepsilon_H = \mathcal{O}(\|x^0 - x^*\|)$ .  $\diamond$

**Satz 5.12. (Lineare Konvergenz des Newton-ähnlichen Verfahrens)**

Es sei (5.5) erfüllt. Dann existieren Konstanten  $K \in (0, 1)$  und  $\delta > 0$ , so dass für  $x^0 \in B(x^*, \delta)$  gilt: Die Folge der Iterierten  $(x^k)_{k \in \mathbb{N}}$ , erzeugt durch das Verfahren (5.7), konvergiert linear gegen  $x^*$  und man hat

$$\|x^{k+1} - x^*\| \leq K \|x^k - x^*\|.$$

**Beweis.**

Sei  $\delta > 0$  so gewählt, dass Satz 5.10 angewendet werden kann. Mit den Bedingungen (5.8) folgt:

$$\begin{aligned} \|x^{k+1} - x^*\| &\leq \overline{K} (\|x^k - x^*\|^2 + \gamma (\|x^0 - x^*\| + \|x^* - x^k\|)) \|x^k - x^*\| \\ &= \overline{K} \underbrace{\|x^k - x^*\|}_{\leq \delta} (1 + \gamma) + \gamma \underbrace{\|x^0 - x^*\|}_{\leq \delta} \|x^k - x^*\| \leq \overline{K} (1 + 2\gamma) \delta \|x^k - x^*\|. \end{aligned}$$

Um Konvergenz zu garantieren, verkleinere  $\delta$ , sodass  $\overline{K}(1 + 2\gamma)\delta < 1$  erfüllt ist.  $\square$

**5.3. Das inexakte Newton-Verfahren**

Während beim klassischen Newtonverfahren das Update  $x_+ = x_a + d_a$  in jedem einzelnen Schritt das (exakte) Lösen des Gleichungssystems  $\nabla^2 f(x_a) d_a = -\nabla f(x_a)$  erfordert, verwendet man beim **inexakten Newton-Verfahren** eine approximative Richtung  $\tilde{d}_a$ , die mit geringerem Aufwand berechnet werden kann und für ein  $\eta_a > 0$  die folgende Bedingung erfüllt:

$$\|\nabla^2 f(x_a) \tilde{d}_a + \nabla f(x_a)\| \leq \eta_a \|\nabla f(x_a)\|. \quad (5.9)$$

Wir wissen, dass  $\nabla^2 f(x_a)$  positiv definit ist für  $x_a$  nahe  $x^*$ . Daher eignet sich zum Beispiel das **CG-Verfahren** (Verfahren der konjugierten Gradienten) zur iterativen Lösung von  $\nabla^2 f(x_a) \tilde{d}_a = -\nabla f(x_a)$ . Man spricht dann vom **Newton-CG-Verfahren**.

**Satz 5.13. (Konvergenzverhalten des inexakten Newton-Verfahrens)**

1. Es sei (5.5) erfüllt. Dann existieren eine Konstante  $K \geq 0$  und ein Radius  $\delta > 0$ , so dass für alle Punkte  $x_a \in B(x^*, \delta)$  mit zugehöriger Richtung  $\tilde{d}$ , welche (5.9) erfüllt, und Update  $x_+ = x_a + \tilde{d}$  gilt:

$$\|x_+ - x^*\| \leq K (\|x_a - x^*\| + \eta_a) \|x_a - x^*\|.$$

2. Sei Voraussetzung (5.5) erfüllt. Dann existieren  $\delta > 0$  und  $\bar{\eta} \geq 0$ , sodass die inexakte Newton-Iteration  $x^{k+1} = x^k + \tilde{d}^k$  mit

$$\|\nabla^2 f(x^k) \tilde{d}^k + \nabla f(x^k)\| \leq \eta_k \|\nabla f(x^k)\| \quad \text{und} \quad (\eta_k)_{k \in \mathbb{N}} \subseteq [0, \bar{\eta}]$$

linear gegen  $x^*$  konvergiert.

Weiterhin gelten:

a) Falls  $\eta_k \xrightarrow{k \rightarrow \infty} 0$ , dann ist die Konvergenz **superlinear**, d.h.  $\lim_{k \rightarrow \infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 1$ .

b) Falls  $\eta_k \leq K_\eta \|\nabla f(x^k)\|$  für ein  $K_\eta > 0$  erfüllt ist, so ist die Konvergenz sogar quadratisch.

**5.3.1. Das Newton-CG-Verfahren**

Bezeichne  $D_h^2 f(x; d)$  eine hinreichend genaue Approximation von  $\nabla^2 f(x)d$ , etwa eine Diskretisierung der zweiten Ableitung von  $f$  im Punkt  $x$  in Richtung  $d$ . Wir betrachten im Folgenden eine vorkonditionierte CG-Methode, welche mit einer Fehlermeldung abbricht, falls  $\nabla^2 f(x)$  in Richtung  $d$  degeneriert, d.h.  $d^T \nabla^2 f(x)d = 0$  gilt, oder  $d$  eine Richtung mit negativer Krümmung ist, d.h.  $d^T \nabla^2 f(x)d < 0$  erfüllt ist.

Wir werden im Kontext von Trust-Region-Verfahren sehen, wie im Fall negativer Krümmungen verfahren werden kann.

**Algorithmus 5.14 (Newton-CG-Verfahren)**

**Eingabe:** Startpunkt  $x^0 \in \mathbb{R}^n$ , Toleranzen  $\tau_{\text{rel}}, \tau_{\text{abs}} > 0$ .

```

1: Setze  $r_0 = \|\nabla f(x^0)\|$  und  $k = 0$ .
2: while  $\|\nabla f(x^k)\| > \tau_{\text{rel}} r_0 + \tau_{\text{abs}}$  do
3:   Wähle Parameter  $\eta_k > 0$  und positiv definite Matrix  $W^k \in \mathcal{S}_n$ .
4:   Setze  $d^0 = 0$ ,  $r^0 = \nabla f(x^k)$ ,  $p^0 = -(W^k)^{-1} r^0$  und  $l = 0$ .
5:   while  $\|r^l\| > \eta_k \|\nabla f(x^k)\|$  do
6:     Bestimme  $w^l = D_h^2 f(x, p^l)$ .
7:     if  $(p^l)^T w^l = 0$  then
8:       return mit Fehlermeldung Indefinitheit.
9:     else if  $(p^l)^T w^l < 0$  then
10:      return mit Fehlermeldung Negative Krümmung.
11:     else
12:       Setze  $\alpha_l = \frac{(r^l)^T (W^k)^{-1} r^l}{(p^l)^T w^l}$ ,  $d^{l+1} = d^l + \alpha_l p^l$  und  $r^{l+1} = r^l + \alpha_l w^l$ .
13:       Setze  $\beta_{l+1} = \frac{(r^{l+1})^T (W^k)^{-1} r^{l+1}}{(r^l)^T (W^k)^{-1} r^l}$ ,  $p^{l+1} = -(W^k)^{-1} r^{l+1} + \beta_{l+1} p^l$  und  $l = l + 1$ .
14:     end if
15:   end while
16:   Bestimme Update  $x^{k+1} = x^k + d^k$ .
17:   if  $f(x^{k+1}) > f(x^k)$  then
18:     return mit Fehlermeldung Kein Abstieg.
19:   end if
20: end while

```

**5.4. Das globalisierte Newton-Verfahren****5.4.1. Abstiegsverfahren**

Bisher haben wir im Zusammenhang mit dem Newton-Verfahren *lokale Konvergenzaussagen* hergeleitet; wir mussten dabei stets voraussetzen, dass der Startpunkt  $x^0$  hinreichend nahe beim Minimum  $x^*$  gewählt war. Um auf diese manchmal problematische Forderung verzichten zu können, führen wir *Globalisierungen* ein, welche die Startpunktwahl relaxieren.

Kann man sicherstellen, dass die **Newton-Iterationsmatrix**  $\nabla^2 f(x^k)$  – oder eine entsprechende Approximation für diese – die Bedingung

$$\forall k \in \mathbb{N}, \forall d \in \mathbb{R}^n : c_1 \|d\|^2 \leq d^T \nabla^2 f(x^k) d \leq c_2 \|d\|^2$$

für gewisse Konstanten  $c_1, c_2 > 0$  erfüllt, so ist  $d^k$  als Lösung von  $\nabla^2 f(x^k) d^k = -\nabla f(x^k)$  eine gradientenähnliche Abstiegsrichtung. Eingesetzt in das Allgemeine Abstiegsverfahren 3.5, folgt dann die globale Konvergenz des **globalisierten Newton-Verfahrens**, d.h. der Startpunkt  $x^0$  kann dann beliebig gewählt werden.

**5.4.2. Die Trust-Region-Methode**

Das klassische Newton-Verfahren könnte als Abstiegsverfahren mit gradientenähnlichen Richtungen aufgefasst werden (wobei das Problem der Schrittweitsuche entfällt), sofern sichergestellt wäre, dass alle Komponenten der Matrixfolge  $(\nabla^2 f(x^k))_{k \in \mathbb{N}}$  positiv definit wären. Das Ziel der Trust-Region-Methode besteht darin, das Verfahren des steilsten Abstiegs, welches keine Anforderungen an die zweiten Ableitungen stellt, in geeigneter Weise mit dem Newton-Verfahren, das höhere Konvergenzraten zur Verfügung stellt, zu verbinden. Wir verwenden dabei eine Umgebung, in der wir einem Modell von  $f$  *vertrauen*.

Sei  $m_a$  ein quadratisches Modell von  $f$  um  $x_a$ , gegeben durch

$$m_a(x) = f(x_a) + \nabla f(x_a)^T (x - x_a) + \frac{1}{2} (x - x_a)^T H_a (x - x_a),$$

wobei  $H_a \in \mathcal{S}_n$  eine Approximation oder gleich der Hessematrix  $\nabla^2 f(x_a)$  ist. Weiter sei  $\Delta$  der Radius einer Kugel um  $x_a$ , in welcher wir dem Modell von  $f$  trauen.  $\Delta$  nennt man den **Trust-Region-Radius** und die Kugel  $\mathcal{T}(\Delta) = \{x \in \mathbb{R}^n \mid \|x - x_a\| \leq \Delta\}$  **Trust-Region**. Die nächste Iterierte wird dann als

approximative Minimalstelle von  $m_a$  in  $\mathcal{T}(\Delta)$  gewählt. Das zugehörige **Trust-Region-Hilfsproblem** lautet daher

$$\min m_a(x + d) \quad \text{u.d.N. } \|d\| \leq \Delta. \quad (5.10)$$

**Bemerkung 5.15.**

Eine einfache Idee zur Lösung des Hilfsproblems beruht auf dem Fixieren der Richtung gemäß des Verfahrens des steilsten Abstiegs unter Berücksichtigung des Trust-Regions. Seien  $x_a$  die aktuelle Iterierte und  $\Delta_a$  der aktuelle Trust-Region-Radius. Der Versuchspunkt  $x_V = x_V(t) = x_a - t_a \nabla f(x_a)$  ist dann definiert über die Minimalstelle  $t_a$  von

$$\min_{t \geq 0} \psi_a(t) = m_a(x_a - t \nabla f(x_a)) \quad \text{u.d.N. } x_V(t) - x_a \in \mathcal{T}(\Delta_a).$$

Es gelten

$$\begin{aligned} \psi_a(t) &= f(x_a) - t \|\nabla f(x_a)\|^2 + \frac{t^2}{2} \nabla f(x_a)^T H_a \nabla f(x_a), \\ \psi'_a(t) &= -\|\nabla f(x_a)\|^2 + t \nabla f(x_a)^T H_a \nabla f(x_a) \\ \psi''_a(t) &= \nabla f(x_a)^T H_a \nabla f(x_a). \end{aligned}$$

Zur Bestimmung von  $t_a$  unterscheiden wir zwei Fälle (vgl. das Problem mit negativen Krümmungen beim Newton-CG-Verfahren 5.14):

1.  $\nabla f(x_a)^T H_a \nabla f(x_a) \leq 0$ : Offensichtlich wird die Trust-Region-Restriktion aktiv, d.h.  $x_V$  liegt auf dem Rand, da  $H_a$  negativ semidefinit.

$$\Delta_a = \|x_V(t_a) - x_a\| = t_a \|\nabla f(x_a)\| \quad \implies \quad t_a = \frac{\Delta_a}{\|\nabla f(x_a)\|}.$$

2.  $\nabla f(x_a)^T H_a \nabla f(x_a) > 0$ : Dann impliziert die Optimalitätsbedingung  $\psi'(t_a) = 0$ , dass  $t_a$  gegeben ist als

$$t_a = \frac{\|\nabla f(x_a)\|^2}{\nabla f(x_a)^T H_a \nabla f(x_a)}.$$

Sollte mit dieser Schrittweite die Trust-Region verlassen werden, wähle  $t_a$  aus 1.

Zusammenfassend setzen wir

$$t_a = \begin{cases} \frac{\Delta_a}{\|\nabla f(x_a)\|} & \text{falls } \nabla f(x_a)^T H_a \nabla f(x_a) \leq 0 \\ \min\left\{ \frac{\Delta_a}{\|\nabla f(x_a)\|}, \frac{\|\nabla f(x_a)\|^2}{\nabla f(x_a)^T H_a \nabla f(x_a)} \right\} & \text{falls } \nabla f(x_a)^T H_a \nabla f(x_a) > 0 \end{cases}.$$

Die Minimalstelle  $x_V(t_a) := x_a - t_a \nabla f(x_a)$  des quadratischen Modells  $m_a$  in Richtung des negativen Gradienten heißt **Cauchy-Punkt** und wird mit  $x_a^{\text{CP}}$  bezeichnet. Man kann zeigen, dass der Cauchy-Punkt Voraussetzung 5.16 erfüllt. Die Verwendung des Cauchy-Punktes führt zwar zu globaler Konvergenz, aber unter Umständen ist die Konvergenzgeschwindigkeit lokal nur linear. Die sog. **dogleg-Technik** leistet einen „glatten“ Übergang von der Richtung des steilsten Abstiegs zur Newton-Richtung. Lokal liegt dann quadratische Konvergenz vor, wenn  $H^k = \nabla^2 f(x^k)$ ,  $k \in \mathbb{N}$ , gilt.  $\diamond$

Wir bezeichnen die Lösung von (5.10), die **Versuchslösung**, mit  $d_V$  und setzen  $x_V = x_a + d_V$ . Im Wesentlichen überprüft man, ob das quadratische Modell eine „gute“ Approximation von  $f$  in  $\mathcal{T}(\Delta)$  ist. Dazu definiert man

$$\begin{aligned} \text{ared}_a &= f(x_a) - f(x_V) && \text{(tatsächliche Reduktion)} \\ \text{pred}_a &= m_a(x_a) - m_a(x_V). && \text{(erwartete Reduktion)} \end{aligned}$$

Es gilt:

$$\begin{aligned} \text{pred}_a &= m_a(x_a) - m_a(x_V) \\ &= f(x_a) - f(x_a) - \nabla f(x_a)^T (x_V - x_a) - \frac{1}{2} (x_V - x_a)^T H_a (x_V - x_a) \\ &= -\nabla f(x_a)^T (x_V - x_a) - \frac{1}{2} (x_V - x_a)^T H_a (x_V - x_a). \end{aligned}$$

**Voraussetzung 5.16.**

1. Es gibt ein  $\sigma > 0$  mit

$$\text{pred}_a = f(x_a) - m_a(x_V) = m_a(x_a) - m_a(x_V) \geq \sigma \|\nabla f(x_a)\| \min\{\|d_V\|, \|\nabla f(x_a)\|\}$$

2. Es gibt ein  $M > 0$  mit  $\|d_V\| \geq \|\nabla f(x_a)\|/M$  oder  $\|d_V\| = \Delta_a$

◇

Sei  $0 < \eta_1 < 1$  (z.B.  $\eta_1 = 0.1$ ). Wir setzen  $\varrho := \frac{\text{ared}_a}{\text{pred}_a}$  und gehen folgendermaßen vor.

1. Ist die Reduktion von  $f$  (beschrieben durch  $\text{ared}_a$ ) klein im Verhältnis zur Reduktion von  $m_a$  ( $\varrho \leq \eta_1$ ), so schließen wir, dass  $\Delta$  zu groß gewählt ist. Wir verwerfen dann die Lösung  $d_V$  und verringern  $\Delta_a$ . Dann lösen wir (5.10) noch einmal.
2. Im Fall von  $\varrho > \eta_1$  akzeptieren wir die Lösung  $d_V$  und setzen  $x_+ = x_a + d_V$ . Der Trust-Region-Radius für die nächste Iteration wird dann abhängig von  $\varrho_a$  gesetzt:  $\Delta_+ \geq \Delta$ , wenn  $\varrho \approx 1$  gilt, sonst  $\Delta_+ \leq \Delta$ . In jedem Fall wird  $\Delta_+ \geq \Delta_{\min} > 0$  gewählt, wobei  $\Delta_{\min}$  eine von der Iteration unabhängige untere Schranke für den Radius ist.

Wir geben nun einen Algorithmus an (vgl. [8, Algorithmen 14.3]). Dabei benötigen wir die folgende Bedingung, um die Konvergenz des Verfahrens zu garantieren: Es gibt von  $k$  unabhängige Konstanten  $\alpha \in (0, 1]$  und  $\beta \geq 1$  mit

$$\|d^k\| \leq \beta \Delta_k, \quad \text{pred}_k(d^k) = m_k(x^k) - m_k(x^k + d^k) \geq \alpha \text{pred}_k(d_c^k), \quad (5.11)$$

wobei  $d_c^k$  das Problem

$$\min m_k(x + d) \quad \text{u.d.N.} \quad d = -t \nabla f(x^k), \quad t \geq 0, \quad \|d\| \leq \Delta_k$$

löst (vgl. Bemerkung 5.15).

**Algorithmus 5.17 (Trust-Region-Verfahren)**

**Eingabe:**  $\alpha \in (0, 1], \beta \geq 1, 0 < \eta_1 < \eta_2 < 1, 0 < \gamma_0 < \gamma_1 < 1 < \gamma_2$  und  $\Delta_{\min} > 0$

- 1: Wähle Startwerte  $x^0 \in \mathbb{R}^n$  und  $\Delta^0 \geq \Delta_{\min}$
- 2: **for**  $k = 0, 1, \dots$  **do**
- 3:   **if**  $\nabla f(x^k) = 0$  **then**
- 4:     **stop** und erhalte Lösung  $\bar{x} := x^k \in \mathbb{R}^n$
- 5:   **else**
- 6:     Bestimme eine Lösung  $d^k$  des TR-Hilfsproblems, welche (5.11) erfüllt.
- 7:     Berechne  $\varrho = \frac{f(x^k) - f(x^k + d^k)}{m_k(x^k) - m_k(x^k + d^k)}$
- 8:     **if**  $\varrho_k > \eta_1$  **then**
- 9:        $x^{k+1} = x^k + d^k$
- 10:    **else**
- 11:       $x^{k+1} = x^k$
- 12:    **end if**
- 13:    Bestimme  $\Delta_{k+1}$  gemäß Algorithmus 5.18
- 14:   **end if**
- 15: **end for**

Der folgende globale Konvergenzsatz wird in [8, Sätze 4.10 und 4.11] bewiesen.

**Satz 5.19.**

Seien  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig differenzierbar und nach unten beschränkt. Es gebe eine Konstante  $C > 0$ , so dass  $\|H^k\| \leq C$  für alle  $k \in \mathbb{N}$ . Dann terminiert Algorithmus 5.17 nach endlich vielen Schritten oder jeder Häufungspunkt von  $(x^k)_{k \in \mathbb{N}}$  ist ein stationärer Punkt von  $f$ . Ist weiter  $\nabla f$  gleichmäßig stetig auf  $\Omega \subset \mathbb{R}^n$  mit  $(x^k)_{k \in \mathbb{N}} \subset \Omega$ , so folgt  $\lim_{k \rightarrow \infty} \|\nabla f(x^k)\| = 0$ .

**Algorithmus 5.18** (Trust-Region-Verfahren-Update)**Eingabe:**  $\eta_1, \eta_2, \gamma_0, \gamma_1, \gamma_2, \varrho_k, \Delta_{\min}$  aus Algorithmus 5.17

- 1: **if**  $\varrho_k \leq \eta_1$  **then**
- 2:    $\Delta_{k+1} \in [\max(\Delta_{\min}, \gamma_0 \Delta_k), \max(\Delta_{\min}, \gamma_1 \Delta_k)]$
- 3: **else if**  $\varrho_k \leq (\eta_1, \eta_2]$  **then**
- 4:    $\Delta_{k+1} \in [\max(\Delta_{\min}, \gamma_1 \Delta_k), \max(\Delta_{\min}, \Delta_k)]$
- 5: **else if**  $\varrho_k > \eta_2$  **then**
- 6:    $\Delta_{k+1} \in [\max(\Delta_{\min}, \Delta_k), \max(\Delta_{\min}, \gamma_2 \Delta_k)]$
- 7: **end if**

**5.5. Quasi-Newton-Verfahren**

Wir haben gesehen, dass das (klassische) Newton-Verfahren eine Reihe von Nachteilen hat:

- Zweite Ableitungen werden benötigt (Rechenaufwand).
- Positive Definitheit muss gesichert werden.
- $\mathcal{O}(n^3)$  Multiplikationen (Lösung des linearen Systems mit direktem Verfahren).

**Quasi-Newton-Verfahren** gleichen diese Nachteile teilweise aus:

- Zweite Ableitungen werden durch erste Ableitungen approximiert.
- Positive Definitheit bleibt bei einigen QN-Verfahren erhalten.
- $\mathcal{O}(n^2)$  Multiplikationen bei QN-Varianten, welche direkt  $\nabla f(x^k)^{-1}$  approximieren.

$H^k$  wird dabei von Iteration zu Iteration aufdatiert. Die allgemeine Grundstruktur ist wie folgt:

1. Wähle eine Richtung  $d^k = -(H^k)^{-1} \nabla f(x^k)$ .
2. Bestimme das Update  $x^{k+1} = x^k + t_k d^k$  mit Schrittweitenstrategie.
3. Verwende  $x^k, x^{k+1}$  und  $H^k$ , um  $H^k$  zu  $H^{k+1}$  aufzudatieren.

**Bemerkung 5.20.** (Aufdatierungsformeln)

Seien  $s_a = t_a d_a = -t_a H_a^{-1} \nabla f(x_a)$ ,  $y_a = \nabla f(x_+) - \nabla f(x_a)$  und  $x_+ = x_a + s_a$ . Dann gilt

$$\begin{aligned} y_a &= \nabla f(x_+) - \nabla f(x_a) \\ &= \nabla f(x_a) + \nabla^2 f(x_a)(x_+ - x_a) + \mathcal{O}(\|x_+ - x_a\|) - \nabla f(x_a) \\ &= \nabla^2 f(x_a) s_a + \mathcal{O}(\|x_+ - x_a\|) + \mathcal{O}(\|s_a\|). \end{aligned}$$

1. Eine nahe liegende Forderung für das Update  $H_+$  von  $H_a$  ist die **QN-Bedingung** oder auch **Sekantenbedingung**

$$H_+ s_a = y_a. \tag{5.12}$$

Ein einfacher Ansatz für  $H_+$  ist dann die **symmetrische Rang 1-Korrektur**  $H_+ = H_a + \alpha u u^T$  für zu bestimmendes  $\alpha \in \mathbb{R}$  und  $u \in \mathbb{R}^n$ . Einsetzen in die QN-Bedingung liefert  $H_+ s_a = H_a s_a + \alpha u(u^T s_a) = y_a$ , d.h.  $u$  ist proportional zu  $y_a - H_a s_a$ . Wähle daher  $u = y_a - H_a s_a$  und berücksichtige die Länge im Parameter  $\alpha$ : Aus  $\alpha(u^T s_a) = 1$  folgt  $\alpha = (y_a^T s_a - s_a^T H_a s_a)^{-1}$ . Also lautet die Update-Formel für  $H$ :

$$H_+ = H_a + \frac{(y_a - H_a s_a)(y_a - H_a s_a)^T}{(y_a - H_a s_a)^T s_a}.$$

Die Rang-1-Korrektur hat aber zwei Nachteile: Die positive Definitheit von  $H$  geht meist verloren und  $y_a - H_a s_a$  liegt nahe bei 0.

2. Flexibler sind **Rang 2-Korrekturen**  $H_+ = H_a + \alpha u u^T + \beta v v^T$ . Eingesetzt in die QN-Bedingung ergibt sich  $H_+ s_a = H_a s_a + \alpha u(u^T s_a) + \beta v(v^T s_a) = y_a$ . Die Vektoren  $u$  und  $v$  sind dann allerdings nicht mehr eindeutig bestimmt. Es bietet sich an,  $u = y_a$  und  $v = H_a s_a$  zu wählen (**Broyden/Fletcher/Goldfarb/Shanno**). Dann gilt

$$H_a s_a + \alpha y_a y_a^T s_a + \beta (H_a s_a)(H_a s_a)^T s_a = y_a \iff \alpha y_a (y_a^T s_a) + \beta H_a s_a (s_a^T H_a s_a) = y_a - H_a s_a$$

ergeben sich  $\alpha(y_a^T s_a) = 1$  und  $\beta(s_a^T H_a s_a) = -1$ . Die **BFGS-Formel** zum Update der Hessematrix lautet dann

$$H_+ = H_a + \frac{y_a y_a^T}{y_a^T s_a} - \frac{H_a s_a (H_a s_a)^T}{s_a^T H_a s_a}.$$

3. Man kann auch die inverse Hessematrix  $\nabla^2 f(x^k)^{-1}$  durch ein geeignetes  $B^k$  approximieren. Die QN-Bedingung lautet in dem Fall  $B_+ y_a = s_a$ . Verwenden wir die symmetrische Rang 2-Korrektur mit  $u = s_a$  und  $v = B_a y_a$  (**Davidon/Fletcher/Powell**), dann erhalten wir die (**DFP-Formel**)

$$B_+ = B_a + \frac{s_a s_a^T}{s_a^T y_a} - \frac{(B_a y_a)(B_a y_a)^T}{y_a^T B_a y_a}. \quad \diamond$$

**Lemma 5.21. (Positive Definitheit des BFGS-Updates)**

Seien  $H_a \in \mathcal{S}_n$  positiv definit,  $y_a^T s_a > 0$  und  $H_+$  gemäß der BFGS-Formel bestimmt.

Dann ist  $H_+$  symmetrisch und positiv definit.

**Beweis.**

Aus  $H_a$  positiv definit und  $y_a^T s_a \neq 0$  folgt für alle Vektoren  $z \neq 0$ , dass

$$z^T H_+ z = z^T H_a z + \frac{z^T y_a y_a^T z}{y_a^T s_a} - \frac{z^T (H_a s_a)(H_a s_a)^T z}{s_a^T H_a s_a} = \frac{(z^T y_a)^2}{y_a^T s_a} + z^T H_a z - \frac{(z^T H_a s_a)^2}{s_a^T H_a s_a}$$

erfüllt ist. Da  $H_a$  symmetrisch und positiv definit ist, existiert  $H_a^{1/2}$  mit  $H_a = H_a^{1/2} H_a^{1/2}$ . Damit gilt

$$\begin{aligned} z^T H_a s_a &= z^T H_a^{\frac{1}{2}} H_a^{\frac{1}{2}} s_a \leq \|H_a^{\frac{1}{2}} z\| \|H_a^{\frac{1}{2}} s_a\| \\ \implies (z^T H_a s_a)^2 &\leq \underbrace{\|H_a^{\frac{1}{2}} z\|^2}_{z^T H_a z} \underbrace{\|H_a^{\frac{1}{2}} s_a\|^2}_{s_a^T H_a s_a} \\ \implies z^T H_a z - \frac{(z^T H_a s_a)^2}{s_a^T H_a s_a} &\geq z^T H_a z - z^T H_a z = 0. \end{aligned}$$

Ferner gilt  $z^T H_a z - \frac{(z^T H_a s_a)^2}{s_a^T H_a s_a} = 0$ , wenn  $H_a^{1/2} z$  und  $H_a^{1/2} s_a$  linear abhängig sind. Da  $H_a^{1/2}$  regulär ist, gilt in diesem Fall  $z = \lambda s_a \neq 0$ , d.h.  $\lambda \neq 0$ . Also ist  $z^T y_a = \lambda s_a^T y_a \neq 0$  und wir erhalten

$$z^T H_a z \geq \frac{(z^T y_a)^2}{y_a^T s_a} > 0 \text{ für } H_a^{1/2} z = \lambda H_a^{1/2} s_a \quad \text{und} \quad z^T H_a z > \frac{(z^T y_a)^2}{y_a^T s_a} \geq 0 \text{ sonst.} \quad \square$$

**Bemerkung 5.22.**

Die Bedingung  $y_a^T s_a > 0$  ist realistisch. Für quadratische Probleme von der Form  $f(x) = \frac{1}{2} x^T G x + g^T x + b$  mit symmetrischer, positiv definiter Hessematrix  $G$  gilt die Beziehung

$$y_a^T s_a = (\nabla f(x_+) - \nabla f(x_a))^T (x_+ - x_a) = (x_+ - x_a)^T G (x_+ - x_a) > 0 \quad \text{für } x_+ \neq x_a.$$

Für allgemeine Probleme wird  $y_a^T s_a > 0$  durch Verwendung von Schrittweitenstrategien (Wolfe-Powell) sichergestellt.  $\diamond$

Zunächst einmal stellen wir fest, dass  $(y^k)^T s^k > 0$  durch die WOLFE-POWELL-Regel garantiert ist.

**Lemma 5.24.**

Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  differenzierbar. Ist  $H_k$  positiv definit und genügt die Schrittweite  $t_k > 0$  der WOLFE-POWELL-Bedingung, so gilt  $(y^k)^T s^k > 0$ .

**Algorithmus 5.23** (Globalisiertes BFGS-Verfahren)

**Eingabe:**  $\alpha \in (0, \frac{1}{2})$ ,  $\rho \in (\alpha, 1)$ ,  $x^0 \in \mathbb{R}^n$  und  $H_0 \in \mathbb{R}^{n \times n}$  symmetrisch, positiv definit.

- 1: **if**  $\nabla f(x^0) = 0$  **then**
- 2:     **stop**
- 3: **else**
- 4:     **for**  $k = 0, 1, \dots$  **do**
- 5:         Berechne die Suchrichtung  $d^k = H_k^{-1} \nabla f(x^k)$
- 6:         Bestimme die Schrittweite  $t_k > 0$  mit Hilfe der WOLFE-POWELL-Regel.
- 7:         Setze  $x^{k+1} = x^k + t_k d^k$
- 8:         **if**  $\nabla f(x^{k+1}) = 0$  **then**
- 9:             **stop**
- 10:         **end if**
- 11:         Setze  $s^k = x^{k+1} - x^k$ ,  $y^k = \nabla f(x^{k+1}) - \nabla f(x^k)$  und berechne

$$H_{k+1} = H_k + \frac{y^k (y^k)^T}{(y^k)^T s^k} - \frac{H_k s^k (H_k s^k)^T}{(s^k)^T H_k s^k}$$

- 12:     **end for**
- 13: **end if**

**Beweis.**

Aus der WOLFE-POWELL-Regel und  $d^k = -(H^k)^{-1} \nabla f(x^k)$  folgt

$$\begin{aligned} (y^k)^T s^k &= (y^k)^T (t_k d^k) = t_k (\nabla f(x^{k+1})^T d^k - \nabla f(x^k)^T d^k) \\ &\geq t_k (\rho \nabla f(x^k)^T d^k - \nabla f(x^k)^T d^k) = t_k \underbrace{(1 - \rho)}_{>0} \underbrace{\nabla f(x^k)^T H_k^{-1} \nabla f(x^k)}_{>0} > 0 \end{aligned}$$

was zu zeigen war. □

**Satz 5.25.**

$x^*$  erfülle die hinreichende Bedingung zweiter Ordnung. Erzeugt Algorithmus 5.23 eine Folge  $(x^k)_{k \in \mathbb{N}}$ , die gegen  $x^*$  konvergiert, und gilt  $\lim_{k \rightarrow \infty} \|H_k - H_{k+1}\| = 0$ , so erfüllen die  $H_k$  die DENNIS-MORE-Bedingung und  $(x^k)$  konvergiert  $q$ -superlinear gegen  $x^*$ .

**Beweis.**

TAYLOREntwicklung und die Quasi-NEWTON-Bedingung  $H_{k+1} d^k = y^k$  ergeben

$$\begin{aligned} \|(H_k - \nabla^2 f(x^2)) d^k\| &\leq \|(H_k - H_{k+1}) d^k\| + \|(H_{k+1} - \nabla^2 f(x^k)) d^k\| \\ &\leq \underbrace{\|H_k - H_{k+1}\|}_{\rightarrow 0 \text{ (} k \rightarrow \infty)} \|d^k\| + \|\nabla f(x^{k+1}) - \nabla f(x^k) - \nabla^2 f(x^k) d^k\| \\ &= o(\|d^k\|). \end{aligned}$$

Damit folgt die Aussage aus der Dennis-Moré-Bedingung. □

**Bemerkung 5.26.**

Ist  $\|H_0 - \nabla^2 f(x^k)\|$  hinreichend klein, so kann die Konvergenz von Algorithmus 5.23 für Startwerte, die ebenfalls hinreichend nahe an  $x^*$  liegen, gezeigt werden. ◇

## Index

- Abstiegsrichtung, 11
- Abstiegsverfahren, 11
- Algorithmus
  - Abstiegs-, 12
  - Armijo-, 13
  - Grad-, 14
  - Newton-CG-, 22
  - Trust-Region-, 24
- Armijo-Regel, 13
- Bedingung erster Ordnung, 5
- Bedingung zweiter Ordnung, 5, 6
- BFGS-Formel, 26
- Cauchy-Punkt, 25
- CG-Verfahren, 22
- Definitheit der Hessematrix, 5
- DFP-Formel, 26
- dogleg-Technik, 25
- Fehlerfunktion, 20
- Fehlerschranke
  - absolute, 20
  - relative, 20
- globalisiertes Newton-Verfahren, 23
- Gradient, 4
- gradientenähnliche Richtung, 13
- gradientenähnliches Verfahren, 11
- Gradientenverfahren, 11, 17
- Hessematrix, 5
- Konvergenzordnung, superlineare, 22
- konvex
  - gleichmäßig, 7
  - strikt, 7
- konvexe Funktion, 7
- konvexe Menge, 7
- Maschinen-Epsilon, 21
- Maximalstelle
  - globale, 3
  - lokale, 3
  - strikte globale, 3
  - strikte lokale, 3
- Minimalstelle
  - globale, 3
  - lokale, 3
  - strikte globale, 3
  - strikte lokale, 3
- Minimum
  - globales, 3
  - lokales, 3
  - striktes globales, 3
  - striktes lokales, 3
- Modell, 15
- Modell, polynomiales, 15
- Modul, 7
- Nebenbedingung, 3
- Newton-CG-Verfahren, 22
- Newton-Iterationsmatrix, 23
- Newton-Schritt, 19
- Newton-Verfahren, 19
  - inexaktes, 21
- Optimalitätskriterien, 5
- Optimierungsproblem
  - diskretes, 3
  - endlichdimensionales, 3
  - kombinatorisches, 3
  - nicht-differenzierbares, 3
  - restringiertes, 3
  - stetiges, 3
  - unrestringiertes, 3
- QN-Bedingung, 26
- QN-Verfahren, 26
- Rang 1-Korrektur, symmetrische, 26
- Rang 2-Korrekturen, 26
- Reduktion
  - erwartete, 23
  - tatsächliche, 23
- Restriktion
  - Ganzzahligkeits-, 3
  - Gleichungs-, 3
  - Ungleichungs-, 3
- Sattelpunkt, 5
- Schrittweite, 13
- Schrittweitenstrategie, 11
- Schrittweitenstrategien, 13
- Sekantenbedingung, 26
- Spektralnorm, 5
- stationärer Punkt, 4
- steilster Abstieg, 11
- Trust-Region, 23
- Trust-Region-Hilfsproblem, 23
- Trust-Region-Radius, 23
- Versuchslösung, 23
- Vorwärtsdifferenz, 20
- Winkelbedingung, 12
- Wolfe-Powell-Regel, 15
- Wolfe-Powell-Regel, strenge, 16
- zentrale Differenz, 21
- Zielfunktion, 3
- Zulässigkeitsbereich, 3

## Literaturverzeichnis

- [1] Boyd, S. & Vandenberghe, L.: *Convex Optimization*. Cambridge University Press, 2004.
- [2] Geletu, A.: *Solving Optimization Problems using the Matlab Optimization Toolbox*. Vorlesungsskript Technische Universität Ilmenau, 2007.
- [3] Harbrecht, H.: *Nichtlineare Optimierung*. Vorlesungsskript Universität Stuttgart, 2011.
- [4] Hintermüller, M.: *Nonlinear Optimization*. Vorlesungsskript Universität Berlin, 2010.
- [5] Kelley, C. T.: *Iterative Methods for Optimization*. SIAM Frontiers in Applied Mathematics, 1999.
- [6] Luik, E.: *Numerik I*. Vorlesungsskript Universität Konstanz, 2010.
- [7] Nocedal, J. & Wright, S. J.: *Numerical Optimization*. Springer Series in Operations Research, 1999.
- [8] Ulbrich, M. & Ulbrich, S.: *Nichtlineare Optimierung*. Birkhäuser Verlag, 2012.
- [9] Volkwein, S.: *Numerische Verfahren der restringierten Optimierung*. Vorlesungsskript Universität Graz, 2009.
- [10] Werner, J.: *Optimierung*. Vorlesungsskript Universität Hamburg, 2008.